
A thesis submitted in fulfillment of the requirements for the award of the degree of

Doctor of Philosophy

from

University of Wollongong

by

Riccardo Biondini

BMath. (Hons) & MSci. (Hons), University of Wollongong

Department of Accounting and Finance

2003
In accordance with the rules and regulations of the University of Wollongong, I hereby state that the work described herein is my own original work except where due references are made, and has not been submitted for a degree at any other university or institution.

The research within this thesis has been published in two papers, with a further two papers submitted to journals, awaiting final approval for publishing.

Riccardo Biondini

December, 2003
Acknowledgements

I dedicate this thesis in loving memory of my father and best friend, Sergio, who was always, and still is, there for me. The inspiration of knowing and understanding the feats this great man accomplished during his life fills me with an immense sense of pride and gratitude. None of my degrees, least of all this PhD, would have been possible had it not been for my father’s unfailing support and constant encouragement. A true legend!

I would also like to express my sincere gratitude to my mother, Mariagrazia, for all she has done for me throughout my studies and, indeed, my life. Irrespective of the time I would arrive home, there was always a first-class meal awaiting me. Nothing is ever too much trouble for her. The achievement of obtaining this PhD has as much to do with my mother and my father as it has to do with me. As the English expression goes - “the apple doesn’t fall far from the tree”.

Another very special person to whom I am sincerely indebted is my girlfriend and best buddy, Leyla. She has always had kind words for me and has been extremely supportive and proud of my achievements. I am very appreciative of her faith in me. She understands the importance of study and has motivated me when times were difficult. I thank her from the bottom of my heart for her contribution to the success of my studies. Grazie infinite, Lila.
I would also like to thank my supervisors, Associate Professors Michael McCrae and Yan-Xia Lin, for their involvement with this thesis. I feel fortunate to have had the chance to learn so much from their profound knowledge.

Last but not least, I would like to make special mention of my German Shepherd and companion, Jimmy Schumacher. He has certainly never made life in our household dull and is a wonderful loyal comrade. “Are you moonies, Jimmy? Are you moonies?”.

Dedico questa tesi in memoria di mio padre, Sergio. Sei grande Pa!

Riccardo Biondini
Tra il dire e il fare, c’è di mezzo il mare
Contents

1 Introduction ................................................. 4

1.1 Overview ................................................ 6

1.2 Purpose of Analysis ..................................... 10

1.3 Statement of Problem ................................... 13

1.3.1 Limitations of Previous Research ................. 13

1.3.2 The Consideration of Dynamic Information .... 14

1.3.3 The Consideration of Cointegration Information ... 16

1.4 Motivation of Analysis ................................... 17

1.4.1 Abstracting Dynamic and Cointegration Information ... 18

1.5 Contribution of Research to the Literature ........ 19

1.6 Applications of Analysis ................................ 22

1.7 Methodology in Examining Issues .................... 25

1.8 Dynamic Versus Static Hedging Studies ............ 27

1.8.1 Commodity Hedging ................................ 28

1.8.2 Currency Hedging ................................ 29

1.8.3 Stock Index Hedging ............................. 29
2 The Concept of Hedging 39

2.1 Introduction 39
2.2 Financial Risk 41
2.3 Futures Markets 44
2.4 Hedging Mechanics 47
  2.4.1 Decisions Involved in Hedging 49
  2.4.2 Determination of the Hedge Ratio 51
  2.4.3 Optimal Hedge Ratios 58
  2.4.4 Effectiveness of a Hedge 60
  2.4.5 Hedge Ratio Adjustments 63
  2.4.6 Modelling the Basis 68
2.5 Cross-hedging 74
  2.5.1 Effective Cross-hedging 76
  2.5.2 Foreign Exchange Cross-Hedging 77
2.6 Conclusion 80

3 Conditional Variance Specifications 82

3.1 Introduction 82
3.2 Distributional Characteristics of Returns 84
  3.2.1 Skewness 85
3.2.2 Kurtosis .............................................. 85
3.2.3 Volatility Clustering ................................. 86
3.2.4 Testing for Departures from Normality .............. 88
3.3 ARCH Models ........................................... 91
   3.3.1 The ARCH(q) Model ............................... 93
3.4 GARCH Models .......................................... 95
   3.4.1 The GARCH(1,1) Model ............................ 98
   3.4.2 Multivariate GARCH Models ....................... 100
   3.4.3 Estimation Procedure .............................. 106
3.5 The GARCH-X Model .................................. 109
3.6 Conclusion ............................................. 113

4 Determination of Dynamic Hedging Procedures 115
4.1 Introduction ........................................... 115
4.2 Shortcomings of Conventional Hedging Models ........ 117
4.3 Comparison of Hedging Methods ....................... 120
   4.3.1 Conditional Variance of Different Hedging Strategies .... 121
   4.3.2 Comparisons Between Different Hedging Strategies ....... 122
   4.3.3 Comparison Between Constant Hedge Ratios .............. 125
4.4 Forecasted Hedge Ratios ............................... 126
4.5 Applications of Dynamic Hedging Procedures ........... 132
   4.5.1 Application to Simulated Data ..................... 133
   4.5.2 Application to Financial Data ...................... 138
List of Figures

4.1 Time series plot of the spot returns for Example 1. . . . . . . . . . . 136

4.2 Time series plot of the futures returns for Example 1. . . . . . . . 136

4.3 Time series plot of the hedge ratios obtained via implementation of the
bivariate GARCH model for Example 1. . . . . . . . . . . . . . . . . . . . . . . 137

4.4 Conditional variance of the time-varying (GARCH) hedge minus the conditional variance of the minimum-variance hedge for Example 1. . . . . 139

4.5 Conditional variance of the time-varying (GARCH) hedge minus the conditional variance of the naive hedge for Example 1. . . . . . . . 139

4.6 Conditional variance of the time-varying (GARCH) hedge minus the conditional variance of the forecasted hedge for Example 1. . . . . . . . 140

4.7 Conditional variance of the minimum-variance hedge minus the conditional variance of the naive hedge for Example 1. . . . . . . . . . . . . . . . . 140

4.8 Conditional variance of the minimum-variance hedge minus the conditional variance of the forecasted hedge for Example 1. . . . . . . . . . . . . . . . . . 141

4.9 Conditional variance of the naive hedge minus the conditional variance of the forecasted hedge for Example 1. . . . . . . . . . . . . . . . . . . . . . 141
4.10 Time series plot of the heating oil futures prices for Example 2. . . . . . . . . 143
4.11 Time series plot of the crude oil futures prices for Example 2. . . . . . . . 143
4.12 Time series plot of the heating oil futures returns for Example 2. . . . . . 144
4.13 Time series plot of the crude oil futures returns for Example 2. . . . . . 144
4.14 Time series plot of the hedge ratios obtained via implementation of the
bivariate GARCH model for Example 2. . . . . . . . . . . . . . . . . . . . . 147
4.15 Conditional variance of the time-varying (GARCH) hedge minus the condi-
tional variance of the minimum-variance hedge for Example 2. . . . . . 149
4.16 Conditional variance of the time-varying (GARCH) hedge minus the condi-
tional variance of the naive hedge for Example 2. . . . . . . . . . . . 149
4.17 Conditional variance of the time-varying (GARCH) hedge minus the condi-
tional variance of the forecasted hedge for Example 2. . . . . . . . 150
4.18 Conditional variance of the minimum-variance hedge minus the conditional
variance of the naive hedge for Example 2. . . . . . . . . . . . . . . . . . 150
4.19 Conditional variance of the minimum-variance hedge minus the conditional
variance of the forecasted hedge for Example 2. . . . . . . . . . . . 151
4.20 Conditional variance of the naive hedge minus the conditional variance of
the forecasted hedge for Example 2. . . . . . . . . . . . . . . . . . . . . 151
5.1 Time series plot of the linear combination of the cointegration vector via
the RBC procedure for Example 3. . . . . . . . . . . . . . . . . . . . . . 174
5.2 Time series plot of the linear combination of the cointegration vector via
the RBC procedure for Example 4. . . . . . . . . . . . . . . . . . . . . . 177
5.3 Time series plot of the linear combination of the cointegration vector via the RBC procedure for Example 5. 180

5.4 Cross-plot of the estimate of the cointegration vector via the RBC procedure and the proportion contribution of the smallest eigenvalue for Example 3. 182

5.5 Time series plot of the natural logarithms of the spot US exchange rate (relative to the UK Pound) for Example 6. 185

5.6 Time series plot of the natural logarithms of the futures US exchange rate (relative to the UK Pound) for Example 6. 186

5.7 Time series plot of the linear combination of the cointegration vector via the Johansen Method for Example 6. 187

5.8 Time series plot of the linear combination of the cointegration vector via the RBC procedure for Example 6. 188

5.9 Time series plot of the natural logarithms of the Malaysian Ringgit for Example 7. 189

5.10 Time series plot of the natural logarithms of the Philippine Peso for Example 7. 189

5.11 Time series plot of the natural logarithms of the Thai Baht for Example 7. 190

5.12 Time series plot of the linear combination of the Johansen cointegration vector for Example 7. 191

5.13 Time series plot of the linear combination of the RBC cointegration vector for Example 7. 192

6.1 Time series plot of the spot prices for Example 8. 203
6.2  Time series plot of the futures prices for Example 8. . . . . . . . . . . . . . . . . . . 203
6.3  Time series plot of the spot returns for Example 8. . . . . . . . . . . . . . . . . . 204
6.4  Time series plot of the futures returns for Example 8. . . . . . . . . . . . . . . . . 204
6.5  Time series plot of the hedge ratios obtained via implementation of the
    bivariate GARCH-X model for Example 8. . . . . . . . . . . . . . . . . . . . . . . . . . . 208
6.6  Conditional variance of the time-varying (GARCH-X) hedge minus the
    conditional variance of the minimum-variance hedge for Example 8. . . . . . 208
6.7  Conditional variance of the time-varying (GARCH-X) hedge minus the
    conditional variance of the naive hedge for Example 8. . . . . . . . . . . . . . . . . 209
6.8  Conditional variance of the minimum-variance hedge minus the conditional
    variance of the naive hedge for Example 8. . . . . . . . . . . . . . . . . . . . . . . . . . . 209
6.9  Time series plot of the linear combination of the Johansen cointegration
    vector for Example 9. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 211
## List of Tables

2.1 A possible series of transactions for an investor with a long hedging horizon and the preference to continuously roll the hedge forward. . . . . . . . . . 67

4.1 Summary statistics for the spot and futures returns for Example 1. . . . . 137

4.2 Summary statistics for the heating oil futures returns and crude oil futures returns in Example 2. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 145

4.3 Parameter estimates obtained via implementation of the bivariate GARCH model for Example 2. . . . . . . . . . . . . . . . . . . . . . . . . . . 146

5.1 Estimates of the cointegration vector (for four possible combinations of fitted ARIMA models) for Example 3. . . . . . . . . . . . . . . . . . . . . 174

5.2 Estimates of the cointegration vector (for four possible combinations of fitted ARIMA models) for Example 6. . . . . . . . . . . . . . . . . . . . 187

6.1 Summary statistics for the spot and futures returns for Example 8. . . . . 205

6.2 Parameter estimates resulting from the implementation of the bivariate GARCH and bivariate GARCH-X models respectively for Example 9. . . 212
Abstract

The thesis investigates the problems involved in effectively modelling the (time-varying) basis risk (the risk of large movements in the relationship between the spot price and the derivatives price) and the subsequent calculation of effective hedge ratios. The specific purpose of this research involves the identification of conditional variance models that accommodate the time series properties commonly encountered in many spot and futures return series, by targeting changes in basis volatility over time. The research analyses the problem of incorporating conditional basis variance into the time series model by extending popular specifications to include long-run (cointegration) information.

The analysis investigates whether the omission of cointegration information in the underlying econometric time series model leads to inappropriate modelling of long-run and short-run time series behaviour. Cointegration information, through the squared spread between spot and futures rates, may potentially provide predictive power in modelling volatility of asset returns, volatility that is not captured effectively by the GARCH (1,1) model. Consequences of omitting dynamic adjustments include inadequate modelling of the time series behaviour, sub-optimal decision making and the possible progressive degeneration of such modelling and decision making over time.
The research determines hedging criteria that enables the comparison of the effectiveness of different constant hedge ratios. The comparison of conditional variance measures of various hedge ratios is of interest to the hedger who would like to implement a constant hedge but does not wish to be constrained to choosing either the naive or minimum-variance hedge. An alternative constant hedge ratio to the naive and minimum-variance hedges, termed the “forecasted hedge”, is proposed. The forecasted hedge ratio is based upon the forecasting curves of both the conditional covariance between spot and futures returns and the conditional variance of futures returns, extracting information embedded in the time-varying distribution of spot and futures returns.

Dynamic hedges are compared to constant procedures to determine the conditions under which allowance for conditional variance substantially increases hedging effectiveness. Hedging effectiveness is subsequently defined as the percentage reduction in the variance of the portfolio achieved by implementing a hedged rather than an unhedged position. Where dynamic variance-covariance matrices are effectively modelled, hedge ratios may be constructed that subsequently minimise basis risk.

Another major objective of the research involves the determination of the conditions where periodic re-balancing of the optimal hedge ratio leads to increased hedging effectiveness. The analysis determines criteria that must be triggered in order for an alternative hedging strategy to be better (in terms of risk-reduction) than the strategy currently in effect.

The most important contribution of the thesis is to ensure that in time series modelling of financial series, the basic model adopted is capable of accounting for any
significant short-run and long-run characteristics found to be typical of these series. Concentration is focussed on the calculation of optimal hedge ratios in order to simplify the analysis, by using a very common example of the conditional variance situation. However, the fundamental contribution involves attempting to ensure the basic econometric specification is appropriate in modelling typical time series behaviour so it does not need further adjustment.
Chapter 1

Introduction

The dissertation examines the necessity for time series distributional models that allow for conditional volatility of the basis over the hedge life.\(^1\) The consideration of conditional variance models of basis risk may be necessary in optimal hedge calculation and maintenance. Finance theory says there should be no problem with basis risk in efficient markets. However, theoretical assumptions do not always apply in practice. There may be many situations - even in efficient markets - where the basis is relatively unstable over time due to such factors as conditional variance (for example, market cycles). The major issue in the analysis concerns the observed existence of time-varying volatility between the derivative price and the underlying asset price, and the resultant impact of this time-varying volatility on the effectiveness of the hedge as a means of protecting a position in the underlying asset. Derivatives are securities whose value is derived from the value of an underly-

\(^1\)Abken and Nandi (1996) define volatility as a measure of dispersion of an asset price about its mean level over a fixed time interval.
ing asset. Derivative markets enable traders to insure against the risk arising from adverse price fluctuations (Lafuente, 2001).

The research is located in the time series statistical modelling literature in general and in the cointegration literature in particular. The main analytics of the thesis concern the development and testing of advanced statistical modelling techniques in the overall context of calculating and maintaining effective hedge ratios. The specific issue addressed in the analysis involves the problem of satisfactorily incorporating information regarding basis time-dependent volatility into the statistical modelling of basis behaviour.

The research within this thesis may be generalised to other applications that involve the modelling of the conditional second moments of the first differences of one or more time series. Where two or more series are involved (and are cointegrated), the cointegration relationship may provide added information in regards to the (more effective) modelling of the conditional second moments.

The objective of this introductory chapter is to define the issues under investigation, indicate their context and significance, and outline the overall structure of the investigation. Section 1.2 explains the purpose of the thesis. Section 1.3 states the problem investigated, the limitations of past research and the logic behind the consideration of dynamic techniques such as cointegration in the construction of hedge ratios. Section 1.4 proceeds to define and motivate the specific issues addressed in this research and the context of the subject matter. Section 1.5 discusses the significance of the analysis and its contribution to the existing literature. Section 1.6 discusses the empirical applicability of the research. Section 1.7 identifies the
analytical and empirical methodology implemented in the investigation. Section 1.8 provides a literature review of past studies comparing dynamic and static hedging techniques, discussing the research works of various authors. Section 1.9 highlights the assumptions made in the dissertation and the limitations of the analysis. Section 1.10 outlines the structure of the thesis.

1.1 Overview

The holder of a physical or financial asset is exposed to the possibility of losses due to adverse changes in the spot price - the underlying value of the asset. The exposure to adverse fluctuations in the spot price arises from the random, unpredictable nature of future asset price movements. In efficient markets, the unpredictability is the consequence of two underlying factors:

1. the efficient market hypothesis of Fama (1965), indicating that all available and expected future information is captured in the current price, subsequently reflecting the intrinsic value of the financial asset.²

2. the arrival and impact of future information onto the market is considered to

²Problematically, the efficient market hypothesis is essentially a dual hypothesis. It implies that all current and expected future information is impounded in the current price of the financial asset and the resulting price fairly reflects the intrinsic or fundamental value of the asset (Fama, 1965). In many instances either one or both of these propositions appear to be violated (see Haugen (1999) for an in-depth discussion of these violations and other issues relating to the efficient market hypothesis).
be essentially random.$^3$

If all current (and expected future) information is accurately reflected in existing spot prices and the arrival of future information is random, the future level and volatility of prices (in both level and direction) is essentially unknown and unpredictable. The asset holder is exposed to random, unpredictable changes in asset prices that may subsequently lead to severe losses in investment or wealth value. These losses may be reduced by implementing protective strategies in derivative markets. The essence of such protection lies in the fact that while future movements of spot prices are essentially random, the relationship between spot and derivative prices remains relatively more predictable.

While a wide range of insurance vehicles are available to assist in risk and loss management situations, in the case of a large number of financial assets the most popular method of obtaining protection is by simultaneously undertaking opposite positions in the spot and derivative markets. Any resultant loss sustained from an adverse price movement in one market should, at least partially, be offset by favourable price movements in the other market. In other words, derivative markets offer a method of insurance through facilitating the transfer of risk from the hedger to other participants, for example, speculators.$^4$

---

$^3$Recent research argues against the strict randomness of price movements, even in efficient markets. Taleb (2001) details the non-Gaussian distributional behaviour commonly displayed by asset prices.

$^4$Speculators usually have a position only in the spot or derivatives market, but not in both. Johnson (1960) reformulated the theory of hedging and suggested that hedging and speculative activities are often combined in the actions of a decision maker.
The insurance aspect of hedging occurs through the asset holder exchanging the potentially large risk caused by unpredictable movements in spot prices, for the relatively small risk of movements in the relationship between the spot price and the derivatives price (basis risk). The opening of off-setting positions in any number of derivative markets such as forwards, futures, options and swaps is known as hedging the underlying (or expected future) asset market position. The important aim of hedging techniques is the off-setting of potential wealth losses due to adverse price movements in the underlying asset market. From hereon, to simplify the explanation, futures contracts are used as a representative example of derivative contracts.

Ghosh (1993) defined the objective of hedging as minimisation of the risk associated with a portfolio for a given level of return. This research examines the exposure to possible adverse financial effects of movements in an underlying spot price. Since futures markets are basically spot markets for standardised futures contracts, hedging is based on the idea that spot and futures markets are well-correlated (that is, strongly related) and move together, in what Chan and Lien (2002) term, a “constellation”. Hedging exchanges spot risk, the risk of adverse price movements in the underlying instrument, for basis risk, the risk that futures prices may move out of line with spot prices over time. Spot risk is typically much higher than basis risk since spot prices tend to be more volatile than the basis.

---

5Forwards and swaps are over-the-counter, privately negotiated agreements between two parties. Futures and options are standardised contracts traded on exchange markets.
Hedging is based on the assumption the spot-futures price spread is small and stable over time. Any significant break-down in this assumption may weaken hedging effectiveness. Therefore, the contribution of this research is significant from two perspectives:

1. a theoretical econometric perspective, in terms of improved modelling of spot-derivative price spreads.

2. an empirical perspective, in terms of creating and maintaining hedging effectiveness in actual hedging strategies.

Though hedging is a form of insurance, it seldom provides complete protection. For a futures contract to be an effective hedging vehicle, not only should the futures price correlate well with the spot price, but spot and futures prices should converge to each other as the futures contract nears expiration (Chan and Lien, 2002). However, the basis may be unpredictable. Therefore, hedgers run the risk the basis may move against them going forward. Basis risk must be sufficiently moderate, or otherwise modelled effectively, in order to provide an effective hedge.

Hedging is useful for corporations or investors who have an exposure to price movements in physical/financial assets and who desire to reduce the risk associated with this exposure. Hedging is carried out in the futures market by either selling futures in advance of future spot market sales, or buying futures in advance of future spot market purchases (Chance, 2001). Hedging effectively locks in the forward price

Such a break-down may occur in cross-hedges and/or in situations of market turmoil, such as turning points (Castelino, 1992).
when the hedger buys or sells the futures contract, providing spot and futures returns are perfectly correlated. Hedgers usually aim to set a price level in advance for an asset they later intend to buy or sell. They forgo the opportunity to benefit from favourable price movements, while protecting against unfavourable fluctuations. The futures market position is usually cancelled when the spot transactions have been completed so the hedger no longer holds an outstanding uncovered position.

1.2 Purpose of Analysis

The thesis addresses the problem of satisfactorily incorporating information regarding basis time-dependent volatility into the statistical modelling of basis behaviour that gives expression to the assumptions underlying the main approaches implemented in calculating effective hedge ratios. The main purpose of the research is to identify conditional variance models that can accommodate the time series properties found in many spot and futures series, by targeting changes in basis volatility over time. Comparative effectiveness relative to static hedging models is judged by the measure of hedging effectiveness.

There are a number of time series formulations currently available to model the behaviour of the basis (and asset prices in general) such as autoregressive models and, more specifically, autoregressive conditional heteroscedastic (ARCH) models. Arguably the most effective to date are the generalised autoregressive conditional heteroscedastic (GARCH) family of models. The current analysis tackles the problem of incorporating conditional basis variance into the time series model by ex-
tending the GARCH model to include long-run information made available through cointegration.

The essence of the analysis is to suggest improvements to GARCH type models to ensure that as many of the characteristics regarding short and long-term behaviour of financial time series are included into the basic time series/econometric model. Incorporating time series behaviour into the basic underlying GARCH model is necessary for appropriate modelling and decision making, and allows a greater understanding of the dynamics behind the generation of spot and futures returns and the relationship between these returns. The major downside to the inclusion of added components involves the model becoming more complex.

In this analysis, dynamic hedges are compared to conventional static procedures to determine whether allowance for stochastic movements in the construction of hedge ratios increases hedging effectiveness. The measure of hedging effectiveness is defined as the percentage reduction in the portfolio variance from maintaining a hedged rather than an unhedged position. By effectively modelling the dynamic variance-covariance matrix, hedge ratios may be constructed that minimise basis risk. The hedge ratios examined are both constant and dynamic in nature.

The naive and minimum-variance approaches are static risk management strategies that involve a one-time decision about the best hedge and do not require any adjustment to the hedge ratio once this decision has been taken. As an example, the minimum-variance hedge recognises the correlation between spot and futures prices may be less than perfect and estimates the hedge ratio as the minimum-variance coefficient of a regression of spot returns on futures returns. However, the main
shortcoming of this measure is that it imposes the restriction of a constant joint distribution of spot and futures price changes - this restriction may lead to sub-optimal hedging decisions in periods of high basis volatility.

The analysis introduces an alternative constant hedge ratio to the naive and minimum-variance hedges, termed the “forecasted hedge”. The forecasted hedge takes advantage of the added information intrinsically embedded in the dynamic distribution of spot and futures returns. The forecasted hedge is based upon the forecasting curves of the conditional covariance between spot and futures returns and the conditional variance of the futures returns. Hedging criteria are determined that enables the comparison of various constant hedge ratios. The comparison of conditional variance measures of different hedge ratios is of interest to the hedger who would like to implement a constant hedge but does not wish to be constrained to choosing either the naive or minimum-variance hedge.

The thesis focusses on examining conditional variance models that can accommodate the time series properties found in many spot and futures series by targeting changes in basis volatility over time. In particular, three propositions are investigated:

1. the conditions under which time-varying hedging strategies are superior to static hedging strategies, in terms of greater risk-reduction.

2. the conditions under which time-varying hedging strategies, incorporating cointegration information into the conditional variance and covariance equations, are dominant over dynamic hedging strategies that ignore cointegration,
in terms of greater risk-reduction.

3. the conditions under which periodic re-balancing of the optimal hedge ratio - on the basis of conditional variance - dominates in terms of spot risk immunisation.

1.3 Statement of Problem

The previous section introduced the concept of hedging, a detailed discussion of which is included in Chapter 2. This section states the issue investigated in the thesis - the problem associated with effective hedging presented by time-varying volatility in the basis at any point in time. Such a problem is shown to exist by noting the limitations of traditional hedging procedures invoked in previous studies. A discussion of the logic behind the consideration of time-varying procedures takes place.

1.3.1 Limitations of Previous Research

The common problems encountered in previous investigations (both theoretical and empirical) into optimal hedging issues include:

1. the implicit assumption the risk in spot and futures markets is constant over time, disregarding the possible dynamic nature of the distribution of spot and futures returns. The minimum-variance hedge ratio is constant, irrespective of when the hedge is undertaken (Brooks, Henry and Persand, 2002).
2. the omission of information pertaining to both the short-run dynamics and likely long-run cointegration between assets (Kavussanos and Nomikos, 2000a, 2000b).

In regards to the first issue, where the conditional covariance between spot and futures returns does change over time, the joint distribution is not constant, but time-varying. Even though basis risk may be minimised, there is still the risk that over time spot and futures prices may not move together, resulting in a possible gain or loss at the termination of the hedge. In such cases, adjustments should be made to the hedge ratio over time. The omission of adjustments may make the hedge less effective.

The second issue - the omission of the cointegration relationship when constructing hedging decisions - may also be a significant shortcoming of studies that ignore this information. Assume a situation where spot and futures prices are known to be cointegrated, the next logical task would involve an analysis of the short-run dynamics over time and investigation into the extent to which the conditional covariance between the two prices varies in the short-run due to the relationship between spot and futures prices not being perfect.

1.3.2 The Consideration of Dynamic Information

If the joint distribution of spot and futures returns changes substantially over time, a constant hedge ratio is not appropriate. Allowance should be made for the possible stochastic nature of the returns. The empirical results indicate the hedging potential
provided by stochastic hedging rules are quite promising. Cecchetti, Cumby and Figlewski (1988) and Kroner and Sultan (1993) both showed that allowing for time-varying hedge ratios results in a greater reduction in risk (than constant hedging procedures provide). Both Cecchetti, Cumby and Figlewski (1988) and Kroner and Sultan (1993) concluded that conditional hedging techniques are appropriate because, as new information is received and digested by the market, the riskiness of each of these assets changes, subsequently resulting in a re-evaluation of the current price of the asset. The resultant hedge ratios provide greater risk-reduction than conventional models, even after the consideration of transaction costs (see Cecchetti, Cumby and Figlewski (1988), Kroner and Sultan (1993) and Section 1.8 of this thesis for an overview of the literature on this subject).

If the spread between spot and futures prices varies sufficiently, time-variation of the basis should be taken into account when constructing hedges. The hedge ratio may be recalculated using the information from the conditional variance and the conditional covariance. If the hedge ratio changes with respect to time and no recalculation is made, the static hedge is potentially less effective than otherwise possible. The hedge ratio is likely to be dynamic in nature when spot and futures contracts do not possess identical characteristics. As a result, the correlation between the two series is likely to be less than perfect. Even though basis risk is minimised, there is still the possibility that over time spot and futures prices may not move together. Such fluctuations in the basis may result in a possible gain or loss at the termination of the hedge. Due to the existence of basis risk, no static hedge ratio can completely eliminate risk (Cecchetti, Cumby and Figlewski, 1988).
1.3.3 The Consideration of Cointegration Information

Where dynamic changes in basis risk are present, cointegration offers a means of incorporating both long and short-term information into optimal hedge calculation models (Kavussanos and Nomikos, 2000a, 2000b). Theoretically, in efficient markets spot and futures prices for identical assets should be cointegrated in the long-run since arbitrage profits are contrary to the efficient market hypothesis. The “no-arbitrage” condition drives the relationship between spot and futures prices (Figlewski, 1984). Even if spot and futures prices are non-stationary, they should not drift too far apart in an efficient market. However, in the short-run volatility may be present in basis risk because short-run fluctuations are likely to alter the conditional covariance between spot and futures returns (Choudhry, 2003).

Cointegration, for the first time, presents a formal definition of long-run equilibrium between prices. Before the notion of cointegration was developed, there was no such formal statistical expression for a stationary long-run relationship. The only manner in which long-run equilibrium was defined statistically involved various metrics and measures of correlation. Correlation analysis is intrinsically a short-run measure valid only for stationary variables (Tarbert, 1998). The stationarity requirement often restricts application to the returns series, detrending the prices and subsequently losing important information in regards to the common stochastic trends between price series. Forbes and Ribogon (2002) indicate that correlation coefficients are upwardly biased in the presence of heteroscedasticity.

Cointegration has brought fundamentally new information in the sense of a
definitive statistical expression, allowing the determination of long-term relationships among non-stationary variables (Roca, 1999). Coupled with error correction models that describe the process by which short-term deviations revert to this long-term equilibrium level, cointegration also provides the tools to quantify both the long-run relationship and the short-run deviations from equilibrium (Harasty and Roulet, 2000).

While cointegration cannot anticipate an individual price level at some future point in time, it can predict this price level given the price of another associated (cointegrated) series. In the presence of cointegration between spot and futures prices, several studies have found the conventional minimum-variance procedure results in an under-hedged position (that is, the hedge ratio is biased downwards) since no account is made for the presence of cointegration, in turn resulting in the mis-specification of the pricing behaviour between spot and futures markets (Ghosh, 1993; Wahab and Lashgari, 1993; Tse, 1995; Brenner and Kroner, 1995; Lien, 1996).

1.4 Motivation of Analysis

The previous section argued that where basis risk is time-dependent, optimal hedge ratio models that take these dynamics into consideration may dominate static models. The problems that exist in past studies on optimal hedging were outlined, notably the omission of hedge ratios that are dynamic in nature, as well as cointegration not being considered in the determination of hedge ratios. The previous section also discussed the theoretical justification for the consideration of cointegration tech-
niques in dynamic hedging strategies. This section discusses both dynamic hedging procedures (formed by ignoring cointegration information) and dynamic hedging techniques (formed by incorporating cointegration information) in formulating and maintaining time-dependent hedge ratios.

1.4.1 Abstracting Dynamic and Cointegration Information

As noted previously, relatively few studies in optimal hedging consider dynamic methods in constructing hedge ratios. Instead, constant hedges are calculated and applied without forward revisions, even though the relationship between spot and futures returns varies through time. Since static procedures ignore new information that arrives at the market (Brooks, Henry and Persand, 2002), they may become less effective. Where this information leads to a change in the relationship between spot and futures returns, consideration should be given to the implementation of conditional hedging models.

Dynamic models allow for the revision of the hedge ratio to incorporate the dynamic nature of the distribution of the returns. Such generalised procedures are preferred to static alternatives as they allow the modelling of both transitory and permanent statistical characteristics of asset returns. In the studies that do consider dynamic methods, the hedge ratios are generally constructed without using any cointegration information. Cointegration enforces long-term equilibrium be-

---

7The term “hedging strategy”, frequently mentioned in the analysis, refers to both constant and dynamic techniques for hedge ratio calculation, as well as the procedures in which these constant (naive, minimum-variance etc.) and dynamic (GARCH, GARCH-X etc.) hedges are calculated.
tween spot and futures prices into the asset price determination model, providing information about the speed of adjustment implied for short-run changes in basis risk. Any cointegration information is usually subsequently incorporated into the model for the conditional mean of both spot and futures returns (Fama and French, 1987; Castelino, 1992; Viswanath, 1993). However, the approach adopted here is to incorporate the cointegration relationship (using the error correction term) into the modelling of the conditional variances and the conditional covariance.

1.5 Contribution of Research to the Literature

This section outlines the significant contribution the thesis makes to the research on dynamic hedging. It has been noted in the literature that dynamic hedging strategies are more effective than static alternatives when the relationship between spot and futures prices is unstable (see Section 1.8). Where basis risk volatility is substantial and where the volatility varies over time, a real problem emerges of how to estimate and maintain as effective a hedge as possible. The general contribution of the analysis is to suggest ways of increasing hedging effectiveness in such situations.

The research makes several important contributions to the literature on conditional hedging models:

1. a theoretical contribution, in terms of improving the underlying time series modelling of basis behaviour by incorporating into existing models additional long-run information that more adequately captures the potential features of basis behaviour.
2. an empirical contribution to both optimal hedge calculation and maintenance.

The extensions made within the analysis also have applications to the improvement of other models such as derivative pricing models.

3. while the improved modelling technique - the GARCH-X formulation - is specifically applied to optimal hedge ratio calculations, the GARCH-X model also has other potential applications. The features of the GARCH-X specification are relevant to derivatives and option price modelling.

The thesis aims to appropriately model the features of spot and futures returns. Some dynamic hedging strategies in the literature incorporate cointegration into the modelling of spot and futures returns (see Ghosh, 1993; Wahab and Lashgari, 1993; Lien and Luo, 1993, 1994; Tse, 1995; Brenner and Kroner, 1995; Lien, 1996) but the incorporation usually occurs within the conditional means. However, dynamic specifications that allow for cointegration information to be incorporated into the modelling of the conditional second moments of asset returns (Lee, 1994) have been found to be of practical benefit. In turn, these conditional variance models have been shown to be effective in hedging underlying spot positions (Kavussanos and Nomikos, 2000a, 2000b). The fundamental difference between including cointegration information through the error correction term into the conditional variance model (as opposed to the conditional mean model) is that the error correction term may be considered a significant explanatory variable in modelling the conditional variance (as opposed to the conditional mean) of spot and futures returns.

Kavussanos and Nomikos (2000a, 2000b) estimate time-varying and constant
hedge ratios, investigating their performance in reducing price risk. Time-varying hedge ratios are generated by a bivariate error correction model with a GARCH error structure, as well as a bivariate error correction model with a GARCH-X error structure (where the square of the error correction term enters the specification of the conditional covariance matrix). The latter specification links the concept of disequilibrium (as proxied by the magnitude of the error correction term) with that of uncertainty (as reflected in the time-varying second moments of spot and futures prices).

The approach adopted in the analysis is similar to Kavussanos and Nomikos (2000a, 2000b). However, there are four important differences. The first distinction involves the proposition of an alternative constant hedge ratio in this analysis. This provides an alternative choice to the static (naive and minimum-variance) hedges commonly adopted in the literature. In this study, the “forecasted hedge” is found to provide a more effective hedge than both the minimum-variance and naive hedges using both simulated and empirical data. The second distinction involves the comparison of different hedging strategies by conditional variance measures, allowing various hedging procedures to be compared to each other at any time. A consequence of this alternative comparison is that the hedger may adopt different hedging strategies at different points in time.

The third distinguishing feature between this research and that of Kavussanos and Nomikos (2000a, 2000b) is that, in the analysis within this thesis, the cointegration relationship - via the square of the error correction term - is only incorporated into the conditional variance, and not into both conditional mean and
conditional variance equations. Finally, comparison between techniques is limited to out-of-sample calculation in the current analysis, whereas Kavussanos and Nomikos (2000a, 2000b) compare techniques both ex-post and ex-ante. The out-of-sample comparisons are indeed the most important since the hedger cannot apply a hedge retrospectively in practice. In reality, a hedger takes all information available to them at a certain point in time and applies a hedge, based on this information, to some future period.

The significance of the research in this thesis lies in its provision of various hedging strategies, allowing the investor to go beyond the usual ordinary least squares or dynamic (without integrating cointegration information) hedging strategies. The research is practical, allowing for the application of dynamic information to apply a hedge that is constant. The development of conditional hedge ratio calculations enables the hedger to constantly compare conditional and constant variance hedges and thus maintain adequate hedging effectiveness, if hedges are formed under constant volatility assumptions.

1.6 Applications of Analysis

The previous section discussed the significance of the research and its contribution to the existing literature. This section conducts an analysis into the practical applicability of the research, enabling an understanding of the situations under which the conditional variance procedures presented here are of practical benefit. The essential objective is to incorporate conditional variance characteristics inherent in
basis risk into optimal hedge ratio calculation, determining criteria for any necessary re-balancing of hedge ratios, to maintain their effectiveness.

The analytical methods presented in the analysis apply in situations where there is a conditional variance associated with basis risk. Of course, the practical benefits of including conditional variance of basis risk in optimal hedge ratio calculation will depend on the nature and size of the conditional variance. Efficient markets classically show little or no conditional variance in the basis. In such instances, spot and futures prices track each other well in high-frequency financial data. However, many other market situations contain the possibility of substantial conditional variance. The variance may be dynamic in commodity markets, in cross-hedging and in small, thin and volatile markets. Even in mature markets, the variance may be conditional at turbulent stages of the market cycle, for example, turning points. These situations may allow substantial increases in hedging effectiveness over static modelling.

Cross-hedging is a risk management tool usually applied where there is not a viable futures or options market in the asset of interest, or where the hedging market does not have sufficient depth or breadth to permit adequate hedging (depth being defined as lack of volume so the hedger might not be able to sell (buy) the futures or option when they so desire, or may only do so at a substantial discount (premium), breadth is defined as the lack of alternative term lengths for futures or options so the hedger is forced to select a term they may not desire, and have to keep rolling the hedge forward).

Volatility in the basis results in a potential mis-specification of the structural
dynamics of the system. The inherent theory and techniques are applicable to the
Asian region, where futures markets are not always existent, let alone efficient.
The feature of time-varying basis risk is thought to be enhanced in such markets
since there is a lack of an identical match (or an acceptable hedging period) to
the underlying spot instrument, forcing investors to consider alternative markets to
hedge their spot exposure.

From a statistical perspective, this research is especially practical where the
relationship between the spot and the futures asset is less than perfect and evolving
through time. The relationship may be less than perfect due to underdeveloped or
non-existent spot and/or futures markets. The relationship between the spot and
futures asset may be dynamic when cross-hedging. In such circumstances the best
that can be done is to hedge in a related asset, the objective obviously being to
select an existing asset whose price is highly correlated with movements in the asset
of interest - by using existing futures contracts that involve similar price fluctuations
with the spot market instrument. Hedging in a related but not identical asset to the
underlying results in a less effective hedge due to the imperfect connection between
spot and futures markets.

The conclusion reached is that the methodology in the analysis is potentially
applicable to a wide range of practical situations, with statistical increases in hedging
effectiveness. In achieving this increased hedging effectiveness, the hedger may adopt
either constant or dynamic hedge ratios.
1.7 Methodology in Examining Issues

The previous sections stated the problem involved in the research, motivated the purpose behind the analysis and discussed its contribution to the literature on dynamic hedging. This section outlines the methods implemented to achieve the research objectives. In achieving increased hedging effectiveness, the hedger may adopt either constant or dynamic hedge ratios.

The dissertation focusses on the potential role of dynamic techniques in effective hedging. The proposition to examine is whether cointegration information, that arises from the knowledge that spot and futures prices move in some sort of long-run equilibrium, may be implemented into the hedging model to help improve the hedging effectiveness of a financial asset. If cointegration information leads to an improvement in the knowledge of how the statistical nature of basis risk operates and what factors contribute to its behaviour, the effectiveness of the hedge may increase.

The research addresses effectiveness in the face of time-conditional volatility of basis risk by attempting to define and characterise the total volatility into permanent and transitory components via the implementation of generalised autoregressive conditional heteroscedastic (GARCH) models (Bollerslev, 1986). The GARCH specification is a time-dependent conditional variance model that estimates the conditional variance and the persistence of shocks to volatility by allowing the second moments of the distribution to change through time (Bollerslev, 1986). The resultant time-varying hedge ratios are then calculated from the estimated variance-covariance
matrix of the model and involve subsequent determination of effective hedge ratios with futures contracts. In such models, the variance of the series is stationary over the long-term (that is, the variance does not increase or decrease ad infinitum, or trend) but does deviate from the stationary process in the short-term.

In regards to cointegration and hedging, the main issue to examine is whether ordinary hedges may be made more effective by incorporating cointegration information. Analysis is conducted into:

1. the major problems involved in attempting to derive effective trading strategies where cross-hedging is necessary, due to the absence of well-developed markets in the spot asset or commodity.

2. the potential contribution of cointegration when cross-hedging. Can cointegration help achieve effective hedges in markets where cross-hedging is a necessity?

There is an analysis into its potential contribution.

The investigation into whether cointegration information is of benefit in constructing effective hedge ratios invokes an extension of the bivariate GARCH model. Cointegration information is augmented to the bivariate GARCH framework (known as the GARCH-X specification) to model the conditional heteroscedasticity of spot and futures returns via the application of a bivariate GARCH-X model (Lee, 1994). Dynamic techniques such as the GARCH-X model allow added information to be obtained in regards to time varying volatilities, assisting in the determination of effective hedges. Analysis and discussion into the contributing factors that dictate whether these techniques are valuable is undertaken.
The cointegration analysis in the thesis uses two software packages:

1. the GiveWin econometric modelling package (Doornik and Hendry, 2001a, 2001b and 2001c).

2. the RATS (Regression Analysis of Time Series) econometric software (Doan, 2000).

GiveWin is a “front end” application where the data is loaded and commands given. The output may be both text and graphical. Specifically in GIVEWIN, PCGIVE is for single equation modelling and PCFIML for multi-equation work (vector autoregressions, cointegration, simultaneous equations etc.). The RATS software package is used to analyse time series and cross-sectional data, developing and estimating econometric models and forecasting. While the primary focus of RATS is on time series data, the package is also used for cross-sectional and panel data sets.

1.8 Dynamic Versus Static Hedging Studies

This section provides a selective review on the comparison between static and dynamic hedging procedures. The purpose of this review is to show dynamic procedures frequently provide evidence of the superiority of dynamic strategies over constant alternatives. Whether dynamic or static techniques are dominant depends primarily upon the relationship between spot and futures prices. The relationship is dependent upon several factors, for example:
1. the type of data examined.

2. the time-frame involved.

3. whether the hedge is terminated before the expiration date of the contract.

4. whether hedging or cross-hedging.

1.8.1 Commodity Hedging

Myers (1991) found strong evidence for rejecting the constant conditional covariance model in favour of the GARCH model in an application to wheat storage hedging. Myers (1991) compared both a static hedging strategy, derived from the calculation of moving sample variances and covariances of past prediction errors of spot and futures prices, and a dynamic technique that involves the application of the hedge ratios obtained from the implementation of the GARCH model. The GARCH formulation provided superior hedging performance over the moving sample variances and covariances model, both ex-post and ex-ante. However, the static hedging model performed equally well as the GARCH model both ex-post and ex-ante, implying the assumption of constant hedge ratios and the subsequent invocation of such a linear regression approach to effective hedge ratio estimation is adequate in this application.

Baillie and Myers (1991) analysed various commodities and concluded that hedge ratios based on time-varying second moments produce greater risk-reduction than constant techniques (both ex-post and ex-ante). The commodities analysed were coffee, corn, cotton, gold, soybean and wheat. The resulting hedge ratios obtained
using the minimum-variance method were compared to those derived from two versions of the GARCH model. Baillie and Myers (1991) reached the opposite conclusion to Myers (1991), in that the hedge ratios are found to be non-stationary and to vary substantially over time as the conditional distribution between spot and futures prices changes.

1.8.2 Currency Hedging

Kroner and Sultan (1993) concluded that allowing for time-varying hedge ratios leads to substantial reduction in the variances of portfolio returns compared to conventional models. They proposed and implemented a dynamic hedging strategy to hedge currency exposure that takes into account both the long-run cointegration relationship between financial assets and the dynamic nature of the distributions of these assets, allowing an investor to re-balance their portfolio only if the potential utility gains from re-balancing offset the losses due to transaction costs. Using both ex-post and ex-ante comparisons, Kroner and Sultan (1993) found the potential risk-reduction obtained from such dynamic techniques is more than enough to offset the transaction costs for most investors.

1.8.3 Stock Index Hedging

Ghosh (1993) analysed stock indices and concluded the hedging strategy derived from the error correction model is more effective in controlling and substantially lowering the risk of adverse price movements by reducing the overall mean square error. Via considering stock index returns in the United States, the minimum-
variance hedge ratio estimates are found to be biased downwards - the hedge ratio is underestimated - due to mis-specification where spot and futures prices are cointegrated and the error correction term is excluded from the regression, resulting in a less effective hedge. The existence of a downward bias is supported by Wahab and Lashgari (1993), Tse (1995), Brenner and Kroner (1995) and Lien (1996), who suggested the hedge ratio in such instances should incorporate the long-run equilibrium relationship as well as short-run dynamics.

Park and Switzer (1995) concluded that estimation of optimal or minimum-risk hedges should adopt time-dependent conditional variance models such as GARCH. They examined heavily-traded stock index futures in the United States and Canadian stock markets and found that dynamic techniques provide an increase in hedging effectiveness over the conventional constant hedging technique for both ex-post and ex-ante periods. In employing a GARCH hedging model, the performance is improved even after allowing for transaction costs. Park and Switzer (1995) stressed that if the joint distribution of stock index spot and futures prices is changing over time, estimating a constant hedge ratio is not appropriate.

Ghosh and Clayton (1996) concluded that hedge ratios formed by incorporating cointegration information through the error correction term - integrating both short-run and long-run information in modelling the data - are superior to those obtained via the traditional minimum-variance method, as evidenced by the likelihood ratio test and out-of-sample forecasts. They noted that prior hedging studies frequently ignore the equilibrium error from the previous period and short-run deviations when constructing hedge ratios. Ghosh and Clayton (1996) extended the traditional price-
change hedge ratio estimation method by applying the theory of cointegration in analysing stock index futures for France, Germany, Japan and the United Kingdom. They developed procedures that hedgers should adopt in order to control the risk of their portfolios more effectively at a lower cost.

Wang and Low (2003) observe the superiority of the GARCH error correction model over conventional hedging techniques in reducing portfolio risks for both international and domestic investors. The dominance of the GARCH model is attributable to the fact that a conditional hedge model appropriately allows investors to utilise the most updated information when making hedging decisions, leading to optimal hedging. In hedging with foreign currency denominated Taiwanese stock index futures, Wang and Low (2003) conclude that, in the presence of international investors’ interest in the underlying market, a foreign currency denominated stock index futures contract is likely to succeed, making the particular futures contract in question popular and extremely liquid.

1.8.4 General Hedging

Koutmos, Kroner and Pericli (1998) noted that dynamic hedge ratios, derived via a multivariate GARCH (MGARCH) model, provide more effective hedges than their static counterparts. The data considered involved mortgage-backed securities in the United States and hedging strategies were compared both on the basis of risk-reduction and expected utility maximisation. The hedge ratios derived via this dynamic technique are found to be ex-post and ex-ante statistically and economically superior to those from conventional models. Koutmos and Pericli (1999) concluded
that dynamic hedging strategies are indeed dominant over static alternatives, providing both risk-reduction and expected utility maximisation, even after transaction costs are incorporated into the analysis. Six different coupons of mortgage-backed securities in the United States were analysed and a dynamic hedging model proposed that included the error correction terms from cointegration relationships in the conditional mean equations, to preserve the long-term equilibrium relationship of the two markets.

Kavussanos and Nomikos (2000a, 2000b) find the GARCH-X specification provides greater risk-reduction than both a simple GARCH model and a constant hedge ratio, though not to the extent evidenced in other markets in the literature. They suggest this may be due in part to the heterogeneous composition of the underlying index. Kavussanos and Nomikos (2000a, 2000b) investigate the performance of both constant and time-varying hedging strategies in reducing freight rate risk in the Baltic Freight Index. The time-varying hedge ratios are constructed using both a bivariate error correction model with a GARCH error structure and an augmented GARCH (GARCH-X) model where the error correction term enters the specification of the conditional variance-covariance matrix.

Like Kavussanos and Nomikos (2000a, 2000b), Choudhry (2003) compares the GARCH and GARCH-X hedging models with static hedging strategies and concludes that employing such time-varying strategies leads to improved hedging performance. Choudhry (2003) examines the effects of the long-run relationship between spot and futures indices on the hedging effectiveness of six stock index futures markets. Various static and dynamic hedging strategies are examined and compared,
investigating both the total sample and the out-of-sample performance of the ratios.

Dynamic techniques have been shown to be superior to naive procedures for a wide range of assets and commodities. These studies justify further examination of dynamic modelling from a perspective of developing better strategies and comparative measures that are conditional upon time. The analysis proposes an alternative constant hedge ratio to the naive and minimum-variance hedge ratios. In addition, mathematical criteria are developed that compare the conditional variances of the returns between different strategies.

1.9 Assumptions and Limitations of Research

The analysis in the thesis is restricted to the natural logarithms of the relevant prices, in addition to the first differences of these natural logarithms (the returns). Noting that security prices follow a log-normal distribution, the natural logarithm price changes are not independently and identically distributed. The conditional density is more accurate than the unconditional density for describing short-term behaviour. The data analysed should have characteristics that:

1. are representative of the general characteristics of the distribution of spot and futures returns over time.

2. are representative of the general characteristics of the co-distribution (covariance) between spot and futures returns over time.

Several limitations to time series modelling impact on the analysis. These limitations include:
1. the significant issue of model selection. The only dynamic variance specifications considered in the thesis are those from the GARCH family. There are many possible alternative specifications that, in practice, may be chosen. Irrespective of the particular framework adopted, the primary aim is the effective modelling of the conditional variance-covariance matrix.

2. the associated costs in applying dynamic hedging strategies. Transaction and other costs are ignored since the sole consideration is in regards to the effectiveness of hedging in minimising risk, and not the efficiency of hedging.\footnote{The cost/benefit issue is important in practical situations but is not considered here.}

However, the objective is to provide investors who may not wish to re-balance frequently with an alternative hedging strategy that only involves re-balancing the hedge when statistical criteria are triggered, mitigating the issue regarding the constant re-balancing of the hedge.

## 1.10 Tasks in the Statistical Analysis

The statistical analysis in the dissertation proceeds as follows: the first task involves setting up a definition of an effective hedge. An effective hedge is defined as a hedge that reduces the risk of the underlying spot position to a level acceptable to the hedger. Therefore, a hedge is regarded as more effective than another hedge if it provides greater risk-reduction. The second task involves forming the hedge ratio. The hedging strategy applied is either constant or dynamic in nature. The third task involves the calculation of the portfolio return at each out-of-sample period. The
portfolio return is defined in the following chapter. The fourth task is to calculate the variance of these portfolio returns for the out-of-sample period. The lower the variance, the more effective the hedge.

The structure of the thesis reflects and contains the analytical tasks necessary to achieve the research objectives. This chapter provided an overview of the research problem, the research objectives, the analytical method and the empirics of the intended investigation. The second chapter highlights the risk management and market context of the research issue. Futures markets are outlined and the reasons why these markets are popular in practice are illustrated, providing the motivation behind the consideration of futures markets in the thesis. The concept of hedging is further discussed and is essential in understanding the purpose of the research in terms of the mechanics behind the implementation and maintenance of effective hedges.

Having reviewed the theory of hedging and the practical tasks in its implementation, the next task involves the use of this foundation analysis to construct and implement different hedging strategies. Understanding the practical limitations of constant hedging techniques, dynamic models for the (conditional) second moments are considered in Chapter 3. The first task concentrates on the distributional characteristics of financial asset returns. While the interest is not specifically on these characteristics, various models that are able to capture these features are outlined. The motivation behind the analysis of dynamic models is that it provides an alternative structure for the variance-covariance matrix of spot and futures returns. The significance of implementing such dynamic specifications in the modelling of
financial time series lies in the fact such models are implemented to investigate the
transitory and the permanent components of the variance of shocks to the system.

The two major elements of the theoretical framework for the subsequent analysis
are identified and described in Chapters 2 and 3. Chapter 2 reviews the mechanics
of hedging. The specific models used to obtain dynamic hedges in this thesis in-
volve the GARCH family of models. Chapter 4 commences the innovative research
accomplished in the thesis. The fourth chapter outlines various hedging strategies
available to the investor that are both constant and dynamic in nature. These tech-
niques are examined later. The mathematical comparison between different hedging
strategies is undertaken by constructing the conditional variance of the return at
each point in time for various hedging procedures. The measures for the conditional
variance of each hedging strategy are compared, allowing the hedger to implement
different hedging strategies and to alternate between these various strategies going
forward. The significance of such a comparative procedure arms the investor with
potentially useful information that may be extracted to effectively hedge.

Since the issue of cost (from implementing and re-balancing the hedge) is not
considered in the analysis, the interest turns to the investigation of whether an alter-
native constant hedge to the minimum-variance and naive hedges may be considered
appropriate. An alternative constant hedge is proposed in Chapter 4 - termed the
forecasted hedge - and is based upon the conditional covariance and conditional
variance forecasts. In other words, conditional information is obtained to produce a
hedge that is constant, in turn providing an alternative to the naive and minimum-
variance hedges. The various static and dynamic hedges are compared using both
simulated and empirical data via determining which procedure minimises the variance of the hedged portfolio.

The major conclusion reached in Chapter 4 is that the dynamic hedging strategy is never less effective than constant alternatives, providing the conditional variance-covariance matrix $H_t$ is effectively modelled. Therefore, where the dynamic hedge ratios do not alter substantially through time, a constant strategy is a useful proxy for these dynamic hedge ratios. In such circumstances, constant hedging strategies are preferred as the costs involved in re-balancing the hedge are likely to outweigh the benefits obtained by decreasing the risk of the hedged portfolio. However, where the relationship between spot and futures returns is time-varying, dynamic hedge ratios are likely to substantially decrease the risk of holding the spot position compared to constant hedging strategies.

The fifth chapter explains the concept of cointegration. The technique of cointegration is used to incorporate long-run information relevant to conditional volatility time series. Standard GARCH models omit this long-run information. The fifth chapter also discusses the applicability of an alternative cointegration procedure, analysing this alternative approach in both a theoretical and empirical framework.

Chapter 6 presents a method of expanding these GARCH-based techniques by incorporating cointegration-based information directly into the GARCH modelling process (Lee, 1994). The uniqueness of the specification of Lee (1994) lies in the fact that cointegration information is usually adopted in modelling the conditional mean equations of asset returns. The GARCH-X model takes advantage of the cointegration information but applies this information to the conditional second
moments of asset returns. The GARCH-X specification is superior to its GARCH counterpart if the conditional variance changes over time as a function of not only past squared deviations from the mean and previous variances, but also as a function of the squared error correction term.

The GARCH-X model is found to be superior to the GARCH model in modelling both spot and futures returns as well as the covariance between spot and futures returns. Therefore, the square of the difference between spot and futures prices is an important predictor variable in the modelling of the conditional variances and the conditional covariance. In turn, dynamic hedges that are based upon this conditional information reduce the variance of the hedged portfolio above and beyond that obtained via the implementation of the GARCH model alone.

The seventh chapter reviews the analysis, summarises the theoretical and empirical results, details the contributions made by the thesis and indicates possible avenues for future research.
Chapter 2

The Concept of Hedging

2.1 Introduction

The introductory chapter highlighted the particular problem under investigation and the limitations of past research. The chapter also discussed the applicability of dynamic techniques in constructing hedge ratios as well as the issues addressed in the thesis. The analysis was shown to be significant and to contribute to the existing literature. The practical applicability of the research and the analytical and empirical methodology were also introduced in Chapter 1. A review was conducted of previous studies in dynamic hedging and its comparison to static hedging. The chapter also listed the assumptions made in the thesis and the limitations of the analysis.

The introduction emphasised that underlying spot prices are unpredictable and this unpredictability exposes asset and portfolio wealth to the possibility of severe losses. Consequently, there is the need to insure against the possibility of such
losses. Off-setting positions in the derivatives market can offer such protection. The thesis deals with the possibility of time-dependent basis volatility by incorporating this characteristic directly into time series modelling underlying optimal hedge ratio calculations.

The purpose of this chapter is to deal with the mechanics of hedging strategies in derivative markets and the role that conditional volatility in basis risk plays in hedging effectiveness. The analysis requires mathematical notation to distinguish between the effectiveness of various types of hedges. This chapter is outlined as follows: Section 2.2 discusses the need for hedging and provides an introduction into the various types of risk present in financial time series - such risk concepts are fundamental to hedging. The distinguishing features of the futures market are outlined in Section 2.3 since the analysis in the thesis is in the context of this type of market. Section 2.4 introduces the mechanics behind the concept of hedging, highlighting the decisions involved in implementing a hedge. This section also highlights the possible methods of calculating the hedge ratio, the concept of the optimal hedge ratio and a measure for examining the effectiveness of the hedge - the research involves both the calculation of the hedge ratio and its corresponding hedging effectiveness measure. Implicit in optimal hedging is the effective modelling of the basis. The efficient modelling of the basis and the situations where basis risk is likely to be large and volatile in practice are also discussed in this section.

Section 2.5 discusses the technique of cross-hedging. Effective hedging may be performed via the implementation of a futures contract whose price movements are highly correlated with the price movements of the underlying asset. Since the disser-
tation contains several examples of foreign exchange hedging, Section 2.5 discusses the difficulties involved in hedging in the presence of thin or non-existent foreign exchange futures markets.

2.2 Financial Risk

Risk is a fundamental concept in effective hedging. The objective of this investigation is the minimisation of a particular component of risk - systematic price risk. The thesis concentrates on the general effectiveness and efficient management of systematic price risk that, due to modelling limitations, cannot be diversified away. Defining risk and its components is the first task in the analysis and the focus of this section.

The concept of risk plays a fundamental role in financial decisions. Risk is generally defined as a measure of the variability or uncertainty of potential future outcomes of an investment (Kimball, 2000), leading to the possibility that financial transactions do not achieve their desired results (Hunt and Terry, 2002). Chance (2001) describes risk management as the practice of identifying the risk level a firm (or investor) desires, determining the risk level currently in effect, and using derivatives and other available financial instruments to adjust the actual risk level to the desired risk level.

Financial risk has two components: systematic risk and non-systematic risk. Systematic risk, also known as market risk, is the risk that relates to the market or economy as a whole and therefore cannot be diversified away simply by holding a
greater variety of securities. Non-systematic risk, on the other hand, can be largely eliminated by diversification within an asset class, and is often called diversifiable risk (Hyman, Dynkin, Mattu and Konstantinovskiy, 2000). Financial business risk is divided into three categories: price, credit and pure risk (Harrington and Niehaus, 2003).

The risk considered in the present analysis is systematic price risk. Harrington and Niehaus (2003) define price risk as the uncertainty over the magnitude of cash flows due to possible changes in input and output prices. Three specific types of price risk are commodity price risk, exchange rate risk and interest rate risk. Commodity price risk arises from possible adverse fluctuations in the prices of commodities. Exchange rate risk, also referred to as currency risk, arises from the possible adverse variation in costs or returns resulting from a change in currency exchange rates (Gastineau and Kritzman, 1999). Gastineau and Kritzman (1999) define interest rate risk as an adverse variation in costs or returns caused by a change in the absolute level of interest rates, in the spread between two rates, in the shape of the yield curve, or in any other interest rate relationship.

Financial risk may have various consequences when either managed incorrectly or disregarded - it may impede the achievement of the operational objectives of the firm. Financial risk may detract from the maximisation of shareholders wealth and/or may result in potential financial loss to the individual or firm. There are various methods

---

9Commodities are physical goods typically produced in agriculture or mining, usually standardised or subject to grading or other classification, that can be the object of commercial transactions (Gastineau and Kritzman, 1999).
for managing financial risks: they involve either purchasing insurance, proactively managing the firm's assets, or hedging (Siems, 1997). The thesis concentrates on the latter. The objective of hedging is the effective and efficient management of risk.\(^{10}\)

In managing risk, it may be undesirable (or impractical) to cover all of the inherent risk. Therefore, the primary objective is not complete risk coverage or complete risk elimination. A certain amount of risk is necessary in order to generate higher expected returns (Donaldson, 2000), with investors seeking to maximise returns subject to a minimum level of risk. Such a trade-off implies that an investor must balance the return contribution of each investment against its contribution to portfolio risk. The choice of a utility function with a higher degree of risk aversion leads to larger hedge ratios, while a less risk averse utility function would yield smaller hedge ratios (Cecchetti, Cumby and Figlewski, 1988). The analysis in the thesis is applicable either to the bona-fide hedger who is interested only in minimising risk, or to a hedger who is relatively risk-intolerant. However, the research may be extended to consider the risk-return utility functions of hedgers.

In summary, the two components of risk are systematic and non-systematic risk. Furthermore, three specific types of risk are commodity price risk, exchange rate risk and interest rate risk. The analysis within the thesis focuses on the minimisation of systematic price risk for a bona-fide or relatively risk-averse hedger.

\(^{10}\)The effectiveness of the hedge relates to the adequate coverage of substantial risk, the efficiency relates to the cost/benefit payoff.


2.3 Futures Markets

The derivative contracts considered in the thesis are those from the futures market, since hedging with futures contracts is one of the most popularly used techniques for managing risk (Goetzmann and Massa, 1999; Lien and Tse, 2000; Faff and McKenzie, 2002; Demirer and Lien, 2003). However, the analysis can be extended to any other derivatives market. This section discusses the operations and functions of the futures market.

A futures contract is a legal commitment to buy or sell a financial asset at a certain future time for a certain price and is normally traded on an exchange. The contracts are covered by formal regulations laid down by the futures exchange, these rules facilitate trading in an orderly manner. Each futures exchange has an affiliated clearinghouse that records all transactions and ensures that all buy and sell trades match, guaranteeing contract performance and supervising the process of delivery for contracts held to maturity (Hull, 2002). Losses incurred by futures traders are guaranteed against possible bankruptcy of the trader by the clearinghouse, requiring daily settlement. An exact delivery date is usually not specified. Instead, the contract refers to a delivery month in which contract settlement must occur. Contracts may be bought or sold in the secondary market at any time prior to expiration. Futures contracts can be terminated in several ways, the most common being by offsetting the futures position and therefore either not taking delivery (if long) or not making delivery (if short) of the asset or commodity.\footnote{To offset an open short futures position before expiry, a seller of a futures contract simply buys back the contract, while a buyer sells the futures contract to close an open long position. The}
The rationale behind the existence of futures markets is the facilitation of hedging by allowing more risk-averse investors to transfer the risk of price changes to speculators willing to bear such risk. Futures markets provide a vehicle for speculators to potentially profit from price fluctuations. Speculators add liquidity to financial and physical markets such as the futures market, making these markets more viable (Mork, 2001). Therefore, futures markets cannot function with the same efficiency without speculators. Since futures markets provide investors with a relatively low-cost way of trading on new information and hedging against adverse price movements (Faff and McKenzie, 2002), these markets provide greater liquidity than their spot market counterparts, making them likely to be the preferred direction of investment for fund managers (Goetzmann and Massa, 1999).  

Where the asset or commodity is delivered at maturity, a cash settlement takes place. Whether the investor makes a profit or loss at maturity depends on how the index has moved in the period since the trading of the futures contract. The full value of the futures contract is not paid or received on establishment of the contract. Instead, both buyer and seller pay an initial margin, which is a small percentage of the value of the contract. The traded price is the basis on which profit or loss is calculated at maturity or on closing out the position, if this takes place before maturity. The variable portion of the contract is the price, determined at the time of the trade in a process known as price discovery (Hodgson, Masih and clearinghouse notes that such a trade cancels out the original position.  

12Ederington and Lee (2002) document the operation and strategies of large traders in the heating oil futures market.
Masih, 2003). The terms of a futures contract are standardised by type, quantity, quality and delivery time and place. In a seminal paper, Silber (1981) concluded that futures contracts whose specifications closely reflect the needs of hedgers are more likely to succeed in practice.\textsuperscript{13}

Futures markets provide two main functions (Chan and Lien, 2002). Such markets

1. provide a structure where traders with different attitudes towards risk may transfer and subsequently manage their risk exposure.

2. enhance the price discovery process by providing additional markets for information.

In terms of the second function, futures markets are valuable in providing real-time market information to a great number of geographically diverse spot market participants (Chan and Lien, 2002), influencing both their behaviour and, in turn, the spot market prices. Traders not physically participating in the futures market are able to estimate the future value of the underlying asset (Hodgson, Masih and Masih, 2003), in turn improving the efficiency of the related spot market by removing any possible arbitrage opportunities between spot and futures markets. The assimilation of prices occurs faster in the futures market than in the spot market due to the lower transaction costs and higher liquidity of the former market (Goetzmann and Massa, 1999).

The practical analysis within the thesis concentrates on hedging via the futures\textsuperscript{13}Contract design is a major factor determining the success of a particular futures contract.
market. This section discussed the futures contract and how such a contract may be traded and terminated. This section also discussed the rationale behind the existence of futures markets, the role of speculators in futures markets as well as the operation and functioning of such markets.

2.4 Hedging Mechanics

Having highlighted the various types of risk and the fundamental characteristics of the futures market, the discussion turns to the pertinent mechanics of hedging that provide the foundations for the analysis. The investigation is concerned with the potential contribution of improved modelling of basis risk and the resultant effect upon improving hedging effectiveness. The purpose of this section is to define hedging and to identify the decisions involved in the implementation of a hedge. Concentration within the thesis is limited to determining the hedge ratio. The section also highlights the possible methods in determining the hedge ratio, defines the concept of the optimal hedge ratio and provides a measure of the efficiency of a hedge. Finally, the technique of adjusting the hedge ratio is discussed, as is the modelling of the basis.

A short (long) hedge is accomplished by taking a long (short) position in the spot market and a short (long) position in the futures market, and may be appropriate when a hedger holds (does not hold) the underlying asset and is concerned about a decrease (increase) in its price. The central theory of hedging is that the price of the underlying financial asset fluctuates in a similar manner to the price of the futures
asset. If an investor is long (short) the spot asset and desires to protect themselves against a possible decrease (increase) in price, a short (long) futures trade may be executed. Selling (buying) the futures contract would substitute for selling (buying) the asset. Since the hedger is short (long) the futures contract they, depending upon how effective the hedge is, may repurchase (resell) the contract at a lower (higher) price, effecting a gain that may at least partially offset the loss in the spot market.\textsuperscript{14} If spot and futures prices are highly correlated, the hedge reduces a substantial portion of the risk.

Although hedging has traditionally been defined as the act of eliminating or substantially lessening the exposure to risk in an existing or contemplated spot position, the reformulation of the theory of hedging (Johnson, 1960) suggested that hedging and speculative activities are often combined in the actions of a decision-maker. Gastineau and Kritzman (1999) note that hedging amongst professional traders involves a position or combination of positions undertaken to profit from an expected change in a spread or relative value (basis arbitrage). In a seminal paper, Ederington (1979) noted the expected basis movement has an impact on the returns to a hedged position. In particular, when there is reason to believe that futures prices are likely to be consistently lower (higher) than spot prices so that a positive (negative) change in the basis can be expected, hedgers can anticipate a reduction (increase) in the returns to the overall portfolio.

\textsuperscript{14}Demirer and Lien (2003) conclude that long hedgers tend to be more active in futures trading (especially in commodity markets) and provide better hedging performance (particularly in currency and stock markets).
The following task in the analysis involves the discussion of the decisions involved in implementing a hedge. The task concentrates on the main issue of the analysis - calculation of the hedge ratio.

2.4.1 Decisions Involved in Hedging

The focus of this analysis is primarily on the method of calculation of the hedge ratio. The number of futures contracts to trade is dependent upon the calculated hedge ratio. This section discusses each of the decisions involved in hedging. The hedging strategy is dependent on the risk situation confronting the hedger. The strategy must be decided so that it will offset any potential loss. A futures position should be established that is profitable if the spot position is losing.

There are several important decisions confronting a hedger in the implementation of a hedge (Chance, 2001). They include, but are not limited to, the determination of:

1. whether a short or long hedge should be implemented.

2. the asset that underlies the futures contract.

3. the number of futures contracts to trade.

4. the termination date of the futures contract.

The most fundamental decision facing the hedger involves whether to go short or long in the futures market - the decision is not always straightforward. The choice between going short or long is critical since the risk in the underlying spot position
is effectively doubled if the hedger implements a short (long) hedge when a long (short) hedge should be chosen. A short or long hedging strategy depends on the type of exposure. Situations that may give rise to the need for hedging include:

1. a current investment.

2. a future liability.

3. a future sale.

The hedger must also decide upon the asset that underlies the futures contract. When the underlying futures asset matches the spot asset to be hedged, this choice is relatively simple. However, when the underlying asset does not have a futures contract based on the same asset, the choice would involve the selection of a futures asset that is highly correlated with the asset to be hedged (Anderson and Danthine, 1981). Such a cross-hedge may in turn be a problem where more than one futures instrument is an appropriate hedging vehicle. In such cases, the choice as to which asset should underly the futures contract may not be clear. Therefore, the effectiveness of the hedge is likely to be lower than otherwise expected via “direct” hedging (Cecchetti, Cumby and Figlewski, 1988). Cross-hedging is a problem when the futures price does not correspond well to the price behaviour of the hedged asset (Aggarwal and DeMaskey, 1997).

Another crucial decision involved in the implementation of a hedge is the calculation of the number of contracts in the futures position. The calculation of the hedge ratio is the focus of the thesis. The consequence of a sub-optimal decision may result in a substantial amount of outstanding residual risk. The number of contracts
is determined by calculating what is known as the hedge ratio. The hedge ratio is defined as the ratio of the number of units of the derivatives asset that are purchased relative to the number of units of the spot asset held, in an attempt to capture the relationship between changes in the spot price to changes in the futures price when the hedge is lifted (Chance, 2001). When calculating the number of contracts to hedge, the optimal number may not be an integer. Since partial contracts are not issued, the hedger either slightly over-hedges or under-hedges.\footnote{Whether the hedger “over-hedges” or “under-hedges” depends upon whether the number of contracts is rolled “up” or “down”.}

Another decision confronting the hedger in selecting an appropriate futures contract involves the choice of the expiration date of the futures contract. In general, an expiration date should closely correspond to, but be after, the month in which the hedge is terminated (Chance, 2001). Where the hedging horizon is relatively long, a shorter expiration date should be chosen and the contract rolled forward as the expiration date approaches, at which time the old position is closed out and a new position opened in the following expiration month. The risk generated from the rolling of the hedge is known as rollover risk (Chance, 2001). Rollover risk increases with the number of rollovers.

\subsection*{2.4.2 Determination of the Hedge Ratio}

The research focusses on determining the hedge ratio(s) to effectively hedge a given spot exposure. The determination of the hedge ratio is perhaps the most convoluted of the decisions confronting the hedger, since there are several possible methods for
calculating such a ratio (Cecchetti, Cumby and Figlewski, 1988). The most common is the minimum-variance hedging technique, where the hedge ratio is calculated by dividing the covariance between spot and futures returns by the variance of futures returns. A variation of the minimum-variance approach involves replacing the covariance and variance in the minimum-variance hedge ratio by the conditional covariance and conditional variance.

A second method for calculating the hedge ratio is known as the price-sensitivity hedge (or the duration-based hedge), developed by Kolb and Chiang (1981). The price-sensitivity hedge ratio is applied to hedges of interest rate sensitive securities and provides the optimal number of futures contracts to hedge against interest rate changes. The price-sensitivity hedge is examined in detail in Hull (2002). There are also various hedges using options that may be implemented to cover the underlying spot exposure such as the delta, theta, vega, gamma and rho hedges. Since the thesis is not concerned with interest sensitive securities or options, the price-sensitivity hedge ratio is not considered in the analysis. Measures such as hedging effectiveness are not valid for the price-sensitivity or option-based hedges.

Perhaps the most elementary method to hedge risk in the same asset as the underlying spot position is by implementing what is known as the naive (or one-to-one) hedging rule (Wang and Low, 2003). The naive hedging technique involves adopting an opposite position (though equivalent in size) in the futures market to the position taken in the spot market. The naive hedge assumes there is a high positive correlation between spot and futures prices and the hedge ratio is accordingly chosen to equal one (full hedging). However, full-hedging is based on
the assumption the price of the futures contract rises (falls) by the same amount the spot price rises (falls). The assumption that the spot and futures prices change by the same magnitude may be considered tenuous in many empirical examples, motivating the implementation of either the minimum-variance hedge - based on the assumption the price of the futures contract rises (falls) by the same proportion that the spot price rises (falls) - or time-varying techniques. The analysis concentrates on the minimum-variance approach as well as time-varying hedge ratios. An alternative constant hedge is also proposed based on time-varying information.

**The Minimum-Variance Hedge**

As mentioned, since the empirical examples do not consider option prices, the price-sensitivity (duration-based) hedge is not applied. Rather, the minimum-variance technique is used to construct hedge ratios. The naive hedge is also adopted for comparative purposes. This section discusses how past research arrived at the minimum-variance hedge and why such a hedging strategy is plausible.

The original framework of Markowitz (1952) involved a two parameter (risk and return) utility maximisation approach, where the investor’s degree of risk aversion plays an important role in the hedging strategy. Working (1953a, 1962) portrayed the exporter as a risk-taker who behaves much like a part-time speculator and focusses on profit maximisation - under this view the exporter adopts a selective (discretionary) hedging strategy by hedging in anticipation of a favourable relative movement between spot and futures prices. The level of hedging effectiveness depends upon the movements of spot and futures prices and on reasonable assumptions.
of such variations in this relationship (Working, 1953b).

The seminal papers of Johnson (1960), Stein (1961) and Ederington (1979) synthesised the traditional portfolio and hedging theory of Working (1953b) to view hedging as an extension of portfolio theory (Markowitz, 1952). Ederington (1979) postulated the objective of hedging in terms of (minimising) the variance of the expected return of the hedged portfolio. If the traditional price-change (returns) hedging model is implemented and the hedger’s objective is solely risk minimisation in the manner of Ederington (1979), the resultant minimum-variance hedge ratio \( h \) is determined by regressing the returns to holding the spot asset, \( \Delta S_t \), on the returns to holding the futures asset, \( \Delta F_t \), that is,

\[
\Delta S_t = a + h \Delta F_t + e_t, \tag{2.1}
\]

where \( a \) and \( h \) are constants and \( \{e_t\} \) is a sequence of independently and identically distributed random variables. The coefficient \( h \) determines the hedge ratio on the assumption that spot returns are not always equal to the returns of futures or options, there are no substantial autocorrelations in both spot and futures returns, and the hedge ratio remains stable over time. As a result of hedging a proportion \( h \) of the underlying spot position, a fraction \( (1 - h) \) of this spot position remains unhedged. Brooks, Henry and Persand (2002) suggest that since risk is measured as the volatility of portfolio returns, the calculation of the hedge ratio via Equation (2.1) is an intuitively plausible strategy.

The calculated hedge ratio may be negative (indicating the investor should increase exposure beyond what is already in the portfolio), or greater than one (in-
dicating the portfolio should have a net short position). However, it is common
to restrict the range of the hedge ratio to be between zero (the non-hedged posi-
tion) and one (the completely hedged position) since the futures price usually has
the same, if not higher, volatility than the spot price (Demirer and Lien, 2003).
The hedge ratio should be no greater than the correlation between spot and futures
prices, and when \( h < 1 \), every investor would find the over-hedged \( h = 1 \) position
dominated by the risk minimising hedge.

Utility Maximisation and Risk Minimisation Hedge Ratios

The hedge ratio is either calculated by maximising a mean-variance expected utility
function or by minimising the risk of a given spot exposure. This section shows
the hedge ratio obtained via the minimum-variance technique, \( h \), is equal to the
optimal hedge ratio for a utility maximising hedger - with quadratic utility function
- if either the coefficient of risk aversion is very high or the futures price follows
a martingale process. The importance of this investigation is that the research
is generally applicable to all hedgers, irrespective of whether they are solely risk
minimisers or utility maximisers.

If the investor’s objective is to maximise expected utility, the mean-variance
expected utility function to be maximised is

\[
EU(\Pi_t) = E(\Pi_t) - \gamma Var(\Pi_t),
\]

where \( \Pi_t = \Delta S_t - h_t \Delta F_t \), and \( E(\Pi_t) \) denotes the expected value of the total portfolio
return at time \( t \), \( \gamma (> 0) \) is a measure of the investor’s level of risk aversion and
\( Var(\Pi_t) \) is the variance of the portfolio return. \( U(\Pi_t) \) may be interpreted as a risk-adjusted rate of return and is dependent upon the investor’s level of risk tolerance, 
\( h_t \) is the hedge ratio applied at time period \( t \) (based upon information available up to and including time period \( t - 1 \)). Where the hedge ratio, \( h \), is constant, the equation to be maximised is

\[
\begin{align*}
max_h E(U(\Pi_t)) &= max_h [E(\Delta S_t) - hE(\Delta F_t) \\
&- \gamma [Var(\Delta S_t) + h^2 Var(\Delta F_t) - 2h Cov(\Delta S_t, \Delta F_t)]] \quad (2.2)
\end{align*}
\]

Two distinct components form the theoretical optimal hedge ratio: the conditional variance minimising hedge ratio and the speculative demand for futures contracts (Koutmos and Pericli, 1999). The optimal hedge ratio is obtained upon differentiating Equation (2.2) with respect to \( h \) and setting this derivative equal to zero, to obtain

\[
h = \frac{-E(\Delta F_t) + 2\gamma Cov(\Delta S_t, \Delta F_t)}{2\gamma Var(\Delta F_t)}
\]

If the futures price follows a martingale process, the above equation reduces to

\[
h = \frac{Cov(\Delta S_t, \Delta F_t)}{Var(\Delta F_t)} \quad (2.3)
\]

The minimum-variance hedge ratio, \( h \), coincides with the optimal hedge ratio for an expected utility maximising agent with quadratic utility function when either the coefficient of risk aversion is extremely high or the futures price follows a martingale process (Moschini and Myers, 2002). The optimal hedge ratio in this instance is a ratio of the unconditional covariance between spot and futures returns to the unconditional variance of futures returns. If spot and futures returns do not
change in the same proportions with respect to each other, a ratio is required of the conditional covariance between spot and futures returns to the conditional variance of the futures returns. In this case, a bona-fide hedger should adopt a time-varying hedge ratio, \( h_t \), equal to

\[
h_t = \frac{\text{Cov}_{t-1}(\Delta S_t, \Delta F_t)}{\text{Var}_{t-1}(\Delta F_t)}
\]  

(2.4)

If the investor desires to maximise their expected utility, the optimal one-period holding of futures at time \( t \) is given by

\[
E_{t-1}U(\Pi_t) = E_{t-1}(\Pi_t) - \gamma \text{Var}_{t-1}(\Pi_t),
\]

where risk is now measured in terms of conditional variances.

The risk minimising hedge ratio therefore changes through time as new information arrives at the market. The conditional model in Equation (2.4) reduces to the conventional model in Equation (2.3) when the joint distribution of spot and futures returns is constant through time (Kroner and Sultan, 1993). Equation (2.4) provides the dynamic correlation between spot and futures returns. According to Hegde (1982), the correlation between spot and futures returns is expected to rise during periods of heightened volatility unless the link between spot and futures markets breaks down, which may happen if the bid-ask spread in the futures market increased to such an extent that trading activity ceased (Watt, 1997).

Watt (1997) concluded that during periods of increased volatility the correlation from a dynamic model exceeds the correlation from the conventional hedge model, indicating that models that assume constant variances and constant correlations
provide adequate hedging performance on average, but fail to provide adequate per-
formance during periods of heightened volatility, by failing to take into account the
dynamic interaction between spot and futures returns. The conventional hedge ratio
typically exceeds the hedge ratio from dynamic hedging models when volatility is
relatively low, but falls below the dynamic hedge ratios during periods of heightened
volatility. Therefore, an investment manager should increase coverage of the cash
position during periods of heightened volatility (Watt, 1997).

In constructing hedge ratios, the calculation depends upon whether the hedger
would like to minimise the risk of the underlying spot position or to maximise the
respective utility function. This section showed the hedge ratios obtained under both
perspectives coincide under certain conditions. Therefore, the research is applicable
to both risk minimisers and profit maximisers.

2.4.3 Optimal Hedge Ratios

Hull (2002) argues that Equation (2.1) provides the optimal hedge ratio in the sense
of minimisation of the variance of the value of the hedger’s position. However,
such unconditional “optimal” hedge ratios are not really “optimal” in the sense of
providing the “most effective” hedge cover going forward, in situations where there
is likely to be substantial conditionality in basis risk. The term “optimal” appears
to mislead as it assumes that a constant hedge, calculated using a prior period, is the
most effective static hedge in a subsequent period. The analysis within concentrates
on finding “effective” hedges and on comparing the effectiveness of various types of
hedges ex-ante. This section discusses the limitations of constant hedge ratios in an
optimality context.

The calculation of the number of futures contracts required for adequate protection relies upon statistical models of the underlying behaviour of spot and futures prices and the relationship between these prices. Many models currently implemented assume the relationship between the spot price of the asset and the corresponding futures price remains stable. However, even in efficient markets, situations frequently arise in which the volatility of the basis is not constant, but time-varying. If ignored, this characteristic can impair the effectiveness of hedge positions.

The assumption of a constant hedge ratio is only valid when basis risk is approximately stable through time. The technique of static hedging is restrictive where the basis is time-varying because it fails to take into consideration all information available to hedgers when they are constructing their hedged positions (Cecchetti, Cumby and Figlewski, 1988; Kroner and Sultan, 1993). The implicit assumption of constant hedge ratios is that the risk in spot and futures markets (via the covariance matrix of spot and futures returns) is constant over time, implying the hedge ratio is the same, regardless of the timing of the hedge.

Constant hedging translates to the variance-covariance matrix being constant throughout the out-of-sample (forecasting) period - equal to the in-sample period’s unconditional variance-covariance matrix. Static hedging approaches therefore cannot generally produce effective hedge ratios in instances where the basis displays time-varying volatility, forcing investors to consider the application of dynamic procedures in managing such risks. In light of these problems, most hedgers re-balance the hedge as circumstances change and new information is received by the market.
In order to determine if a given hedging procedure is effective, a measure of hedging effectiveness must be stated. The next section outlines such a measure, enabling a comparison between various hedging strategies in order to determine the most effective strategy. The necessity for such a hedging effectiveness measure in the context of this analysis stems from the need to critically compare various hedging strategies in the examples later in the thesis.

2.4.4 Effectiveness of a Hedge

A measure of the effectiveness of the hedge is needed to allow the comparison of different hedging strategies. Since the emphasis is on the minimisation of the risk in an underlying spot position, the more spot risk that is eliminated, the more effective the hedge. This section outlines some of the reasons why a hedge may not be effective, as well as providing a measure to determine the effectiveness of a given hedge.

A hedge is effective when potential spot losses are minimised. In such situations, the profit (loss) in the futures market offsets the loss (profit) in the spot market, eliminating all the hedger’s risk. A perfect hedge is only possible where the basis can be predicted with certainty - this may be achievable when the hedging period of the investor matches the futures expiration date and the physical characteristics of the asset or commodity to be hedged exactly match the asset or commodity underlying the futures contract. If either of these features are missing, a perfect hedge is not possible. In such circumstances, risk may be reduced but not eliminated since the relationship between spot and futures prices is non-deterministic, except to the
extent they converge to each other at the settlement date if the futures contract contains appropriate settlement specifications (Chan and Lien, 2002). Most hedges are imperfect for a variety of reasons, namely:

1. the asset to be hedged may not correspond exactly to that underlying the futures contract.

2. the futures expiration date may not coincide with the hedge termination date.

A hedge may not be effective if the hedging instrument is not identical to that underlying the spot market. The spot and futures instruments may not match in relatively underdeveloped (or non-existent) spot and/or futures markets. There might still exist a great deal of basis risk in such markets because the market is thin (and therefore greater premiums might be asked, causing the spread between the two series to be quite large), or the accessibility when one wants to enter the market is either not good or non-existent. Under any of these scenarios, the basis is likely to possess time-varying properties. In this case, risk management has to rely on cross-hedging or indirect hedging (Anderson and Danthine, 1981). The research of Anderson and Danthine (1981) provided a theoretical cross-hedging model from which most empirical analyses are based. While such a hedge may improve upon direct hedging in thin markets, the remaining risk may still be substantial.

A hedge may not be effective if it is lifted prior to the expiration date of a futures contract or the hedging horizon is longer than the life span of a futures contract. Both of these situations may result in the basis at the termination of the hedging period being uncertain. If a hedge is held all the way to expiration, the basis should
converge to zero (and thus the spot price converges with the futures price on the maturity date of the futures contract). A portion of this “convergence” is predictable (Ghosh and Clayton, 1996). More importantly, its predictability is not influenced by the point in time in which the prediction is made. An identical reasoning applies to the uncertainty (variance) surrounding that predictability, which is the true measure of basis risk. Convergence guarantees that basis risk is low close to expiration and higher further away from it, the predictability of which rapidly increases as its reference point approaches contract expiration. Since the basis must converge to zero at expiration, but is uncertain until then, it stands to reason the closer the future is to maturity, the smaller the deviation between spot and futures prices.

While the complete amount of direct and indirect losses due to adverse price movements may be difficult to estimate, the effectiveness of any hedging strategy for a bona-fide hedger may be determined. The effectiveness of such a bona-fide hedge - the sole aim of which is the minimisation of risk - is commonly measured by the percentage reduction in the total risk (variance of the returns) of the hedged position over the unhedged position (Ederington, 1979).\(^\text{16}\) The measure of hedging effectiveness \((HE)\) can be stated mathematically as

\[
HE = 1 - \frac{Var(\Delta S_t - h_t \Delta F_t)}{Var(\Delta S_t)},
\]

or as one minus the ratio of the variance of the hedged portfolio to the variance of the unhedged portfolio. The closer \(HE\) is to 1, the more effective the hedge.

\(^{16}\)Hedging effectiveness measured solely by the percentage reduction of the hedged position over the unhedged position is applicable to a bona-fide hedger whose sole objective is the minimisation of risk.
In contrast to Ederington’s risk minimisation approach to hedging effectiveness, Cecchetti, Cumby and Figlewski (1988) emphasise the role of both risk and utility in measuring the effectiveness of a hedge. Cecchetti, Cumby and Figlewski (1988) adopt a different perspective to Ederington (1979) as to the definition and measurement of the effectiveness of a hedge. They stated that “the optimal hedge ratio is a function of both the risk-return probability curve that is available in the market and the investor’s utility function”. The approach of Cecchetti, Cumby and Figlewski (1988) suggests the implementation of such an optimal strategy involves both estimation of the joint distribution of returns in order to construct the risk-return frontier. Subsequent optimisation of this utility function determines the appropriate hedge ratio. The higher the hedging effectiveness, the better the contract and/or the strategy in reducing the portfolio variance. Given the choice of various derivative products, the hedger should choose the instrument that provides the highest hedging effectiveness (Lien and Tse, 2001), either from the perspective of risk minimisation or utility maximisation.

2.4.5 Hedge Ratio Adjustments

The incorporation of conditional volatility information into optimal hedge ratio calculations through models such as GARCH and GARCH-X raises the probability the original hedge may need to be recalculated at various points over the hedge’s time horizon. The dynamic adjustment is not identical to the traditional adjustments such as “tailing the hedge”, “delta re-balancing” and “rolling the hedge forward”. The third chapter highlights the distinctions between the dynamic adjustment per-
formed in the thesis and other forms of re-balancing. In short, the hedge in this analysis is re-balanced when the conditional variance of the returns for an alternative hedge at a given time is lower than that of the current hedge. The notation and associated procedures involved in this conditional variance comparison are described in the fourth chapter. This section outlines the nature of traditional adjustments to the hedge - necessary to understand the distinction between re-balancing in the traditional sense and the hedge re-balancing performed in this analysis.

Re-balancing the Hedge

Consider a utility maximising investor who only re-balances their portfolio when the financial benefits outweigh the associated costs of re-balancing.\textsuperscript{17} Obviously the increased expected utility arising from the re-balancing of the hedge must offset the incurred transaction costs. Letting $\Delta S_t$ and $\Delta F_t$ denote the changes in spot and futures prices respectively between time $t - 1$ and $t$, the profit equation at time $t$ is given by

$$\Pi_t = \Delta S_t - h_t \Delta F_t, \quad t' < t,$$

(2.5)

where $\Pi_t$ denotes the profit at time $t$ in holding one unit of the spot asset and, for each unit held, adopting a short position in $h_t$ units in the futures market, under the assumption the futures prices are martingales.\textsuperscript{18} Incorporating the transaction

\textsuperscript{17}The re-balancing of the hedge ratio in the examples within this thesis occurs when the conditional variance of the return for an alternative hedge is lower than the conditional variance of the return for the current hedge.

\textsuperscript{18}A martingale is a stochastic process such that its expected value at any future time is the same as its current value - $E[M(t)|(t > 0)] = M(0)$. If futures prices are martingales, the futures
costs involved in re-balancing the hedge, the profit equation at time $t$ becomes

$$\Pi_t = \Delta S_t - h_t \Delta F_t - y,$$

if the investor re-balances the hedge, or

$$\Pi_t = \Delta S_t - h_t' \Delta F_t,$$

if the hedger does not re-balance the hedge, where $h_t'$ is the hedge ratio obtained from the most recent re-balancing of the hedge, and $y$ is the average transaction cost involved in re-balancing one unit of the underlying. The expected return to the portfolio is $-y$ when the hedger re-balances the hedge and zero when the hedger does not re-balance.

Having indicated the unique nature of hedge ratio adjustments implied by conditional variance in the basis, the next task involves the extraction of what is perhaps the most common adjustment performed in practice, that of “tailing” the hedge.

**Tailing the Hedge**

The “tailing” of the hedge is another form of adjustment to the original, optimal hedge ratio, reflecting interest rate effects rather than conditional basis variance. Since this type of re-balancing is not considered in the dissertation, it is outlined only in brief.

Consider a long hedger who has taken a position in the futures market to protect against a possible price increase in the underlying asset or commodity. Since futures price is an unbiased estimate of the expected future spot price. The martingale assumption implies the expected returns from the portfolio are independent, drawn from an identical distribution and unaffected by the number of futures contracts held.
contracts are marked-to-market daily, gains and losses accrue in an interest-bearing margin account. In order to adjust for this daily marked-to-market feature, the hedge position should be reduced to its present value. Adjusting the hedge position in such a manner is commonly referred to as “tailing the hedge” (Hull, 2002). The associated “tailing factor” is due to the marking-to-market mechanism associated with the investor’s position (Lioui and Poncet, 2001).

When the hedge is tailed, there is the possibility of marked-to-market cash outflows that may accumulate to such an extent where the hedger’s assets become depleted. If the hedger is unable to obtain additional funding, the hedge must be liquidated prematurely. To protect themselves against such a possible scenario, many hedgers adjust the number of futures contracts in a hedged position so the present market exposure of the hedge offsets the underlying exposure (Gastineau and Kritzman, 1999), allowing for interest earned or paid from daily settlement. See Figlewski, Landskrone and Silber (1991) for the mathematical calculation of the adjustment factor for both stochastic and deterministic interest rates.

**Rolling the Hedge**

Another type of hedge ratio adjustment is known as “rolling the hedge” forward. Where the desired hedging horizon exceeds the delivery dates of the original futures contracts, the investor must close out one futures position just prior to its delivery month, replacing it with a new position with a later delivery date. The issue is commonly addressed (as mentioned in Example 2) by rolling the hedge over in the month prior to expiration of the futures contract.
Consider an investor who in January wishes to establish a hedge over one year. Futures contracts may be available with settlement dates every month, going out to one year. Since long-dated futures contracts are less liquid than short-dated contracts, the investor may decide to trade six-month futures contracts, rolling the position over just prior to each delivery month. A possible series of transactions in such a strategy are outlined in Table 2.1.

Table 2.1: A possible series of transactions for an investor with a long hedging horizon and the preference to continuously roll the hedge forward.

<table>
<thead>
<tr>
<th>Date</th>
<th>Strategy</th>
<th>Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>Sell June futures</td>
<td>Sell June futures</td>
</tr>
<tr>
<td>May</td>
<td>Close out June position</td>
<td>Sell October futures</td>
</tr>
<tr>
<td>September</td>
<td>Close out October position</td>
<td>Sell February futures</td>
</tr>
<tr>
<td>January</td>
<td>Close out February position</td>
<td></td>
</tr>
</tbody>
</table>

Although contract liquidity is usually an important factor in the choice of futures contract, some additional risk may be generated when rolling the hedge forward. This added risk arises from the mis-pricing of an option or futures contract at the time the old position is closed and the new position opened (Gastineau and Kritzman, 1999). In the case of hedgers with relatively long hedging horizons, forward contracts are more popular than futures contracts since no rollovers occur. The hedger is certain of the final settlement price as this price is pre-determined.

---

19Hedgers should generally be wary of holding positions in the delivery month since the price of the contract may sometimes fluctuate due to erratic volume and open interest behaviour. The general rule of thumb is to select an expiration month that is as close to, but after, the month in which the hedge is terminated.
In the futures market, hedges have to be constantly rolled forward, which leads to rollover risk - basis risk amongst the futures contracts. As a consequence, the precision of the hedge deteriorates with the number of rollovers. Rollover risk is also induced by using short-dated futures to hedge a long-term forward delivery contract. Since there is usually a short period of time when the contract is expected to be rolled forward, the hedger can try to decrease the rollover basis risk by adjusting the time the rollover takes place. Rolling the hedge forward also increases the transaction costs of the hedge.

Having distinguished between traditional hedge ratio adjustments (such as “tailing” or re-balancing) from the form of adjustment implemented in the thesis - adjustment of the hedge ratio implied by conditional variance - the discussion focusses on modelling the basis.

2.4.6 Modelling the Basis

This section discusses the importance of modelling the basis in effective hedging. In particular, the deviations of the basis from its long-run equilibrium level is of interest.\textsuperscript{20} By modelling these short-term fluctuations effectively, the volatility between spot and futures prices is adequately captured, improving hedging effectiveness. In this thesis, the basis is modelled using information that cointegration provides. Cointegration allows the modelling of the long-run equilibrium level as well as short-run deviations from equilibrium and is described both theoretically and mathematically.

\textsuperscript{20}In efficient markets there exists a theoretical relationship between spot and futures prices that dictates these two rates should not drift too far apart - this is the notion of cointegration.
in the fifth chapter. The information that comes from cointegration is included as an explanatory variable (the square of the error correction term) in the modelling of the conditional second moments of asset returns.

The basis fluctuates, but its change is typically less volatile and therefore more predictable than the spot return. Therefore, the hedged position has less risk than the unhedged position.\textsuperscript{21} In general, the larger (smaller) the positive correlation between spot and futures prices, the smaller (larger) the basis risk. Deviations between spot and futures prices are likely not to be large in the long-term due to arbitrageurs beginning to take positions in the two markets, earning excess returns from near-riskless transactions (Figlewski, 1984). Such arbitrage trading therefore eliminates the mis-pricing and brings the deviation between spot and futures prices into line.

Since the basis is defined as the difference between spot and futures prices, and basis risk is the inherent volatility in the basis, simple mathematics would dictate that basis risk decreases as the variance of spot prices and futures prices decreases. Ederington (1979) noted that basis risk as a proportion of total risk decreases with the length of the hedge, thus increasing hedging effectiveness. Relatively longer hedges may allow for resolution of some price uncertainty (“noise”) that would otherwise be reflected in increased basis risk. With lower price uncertainty, hedging effectiveness increases.

One potential source of increased basis risk is cross-hedging - where the asset to be hedged is different from the underlying asset. Increased basis risk may result in

\textsuperscript{21}Basis risk is higher in certain markets (stock market) than in others (foreign exchange market).
the situation where the hedger is uncertain as to when the transaction is likely to be completed and the futures contract may have to be closed out before the expiration date. As noted previously, the general rule is to avoid holding a futures position in its expiration month. Another cause of increased basis risk in the equity market is the non-systematic component of the return on the spot position. Since the index contract is related to the behaviour of the underlying market index, such non-market risk cannot be hedged. Changes in the factors that affect the cost-of-carry and storage, insurance and opportunity costs may contribute to a change in basis risk.

**Relationship Between Spot and Futures Prices**

Effective basis modelling should incorporate the factors that drive the relationship between spot and futures prices, allowing an understanding of the sensitivities of the spot (futures) price to a change in the futures (spot) price. The thesis models the relationship between spot and futures prices through cointegration. The long-run equilibrating relationship between spot and futures prices is found to be an important explanatory variable in modelling the second moments of spot and futures returns. Subsequently, this has an impact on the calculated hedge ratios.

This section outlines the theoretical relationship between spot and futures prices, introducing such concepts as the risk premium, the cost-of-carry and the convenience yield.

---

22Returns to a stock index portfolio may include dividends. However, the index and the index futures only track the capital value of the portfolio. Any risk associated with the dividends of the portfolio becomes basis risk in the hedged position.
The spot price and the futures price are related through a concept known as the risk premium. The risk premium is defined as the return speculators must receive for their willingness to bear risk, thus providing insurance for the hedgers. The risk premium hypothesis states the expected future spot price should be equal to the current futures price plus the expected risk premium. The theory of uncovered interest rate parity states that, in the absence of a risk premium, the futures rate is an unbiased estimate of the future spot rate. Theoretically, the futures price today equals the expected price of the futures contract at expiration because the expected futures price at expiration equals the expected spot price at expiration. In a market with risk premiums, the futures price underestimates the spot price at expiration by the amount of the risk premium.

A contango market exists when the futures price is above the spot price (where no dividends are paid). A contango market is usually due to the cost-of-carry. Where the spot price is greater than the futures price, an inverted market (also known as backwardation) is said to exist. Backwardation is usually due to the convenience yield - the additional return earned by holding a commodity in short supply or a non-pecuniary gain from an asset. Suppose the commodity is in short supply so that current consumption is unusually high relative to the supply of the good - in turn producing an abnormally high spot price. The current tight market conditions discourage individuals from storing the commodity. If the situation is severe enough,

23As a result, a futures contract has an expected return of 0%, a crucial assumption in the determination of the optimal hedge ratio. In such cases, speculators are not rewarded for the risk they bear.
the current spot price may be above the expected future spot price.\textsuperscript{24}

Under the cost-of-carry model, the price of the futures contract is equal to the spot commodity price plus the cost-of-carry, minus the convenience yield. The price of the futures contract may be expressed mathematically as

\[ F_t = S_t + C_{t,T} - k_{t,T}, \]

where \( C_{t,T} \) is the net cost-of-carry that includes the accrued interest expense, storage and insurance costs minus the accrued coupon or dividend yield, and \( k_{t,T} \) is the accrued convenience yield over the period \([t, T]\). Defining the basis at time \( t \), \( b_t \), to be equal to

\[ b_t = S_t - F_t, \]

then

\[ b_t = k_{t,T} - C_{t,T} \]

As time advances, the basis changes. Let \( \Delta b_t \) represent the change in the basis over a small time increment, and let \( \Delta k_{t,T} \) and \( \Delta C_{t,T} \) represent the corresponding changes in the convenience yield and cost-of-carry respectively. Therefore,

\[ \Delta b_t = \Delta k_{t,T} - \Delta C_{t,T} \]

\textsuperscript{24}There may have been a disaster in the coffee crop due to the inclement weather but some farmers have accumulated coffee and have carried it forward from the previous year, thus being able to command a convenience yield. Here the futures price equals the spot price plus the cost-of-carry, minus the convenience yield. The cost-of-carry is the difference between the amount of income generated by an asset and the accrued interest, storage and insurance expenses, minus the accrued coupon or dividend yield.
For financial assets and investment commodities, or for consumable commodities in ample supply, the convenience yield is negligible, and the change in the basis is determined by the change in the cost-of-carry term. The basis for a consumable commodity that is currently in short supply, or anticipated to be in short supply before the delivery date, reflects a convenience yield.

The basis may not change in a continuously predictable manner and may therefore be less predictable than the corresponding change for a commodity with no convenience yield because of the potential for large unanticipated changes in the convenience yield. In particular, unanticipated imbalances between supply and demand can lead to large shifts in the convenience yield, causing the basis to deviate from its predicted level. The basis may unexpectedly expand or contract due to the uncertainty in interest rates. When the basis does move towards zero, a weakening basis is said to occur. Conversely, the basis is said to strengthen when it moves away from zero. Where \( S_t > F_t \), a weakening (strengthening) basis works to the advantage of the short (long) hedger by causing the long (short) position to profit (lose) more than anticipated.

The effective modelling of the basis is an important issue in effective hedging. This section showed the effective modelling of the basis leads to the volatility of spot and futures prices being adequately captured, in turn increasing the effectiveness of the hedging strategy. In particular, the effective modelling of short-term fluctuations of the basis from its long-run equilibrium level is of interest. The concept of

\[ 25 \text{The cost-of-carry term may change in a predictable manner. If interest rates remain constant, the cost-of-carry term } C_{t,T} \text{ smoothly converges to zero.} \]
cointegration allows the modelling of the long-run equilibrium level of the basis as well as short-run deviations from equilibrium. Cointegration information is included as an explanatory variable (the square of the error correction term) in the modelling of the conditional second moments of asset returns. The modelling of the basis is even more important when cross-hedging as the volatility in the basis is likely to be greater. The technique of cross-hedging is outlined in the following section.

### 2.5 Cross-hedging

Cross-hedging may be thought of as a technique to reduce risk when the direct hedging of a position is either inefficient or impossible due to the absence of well-developed derivative markets (Anderson and Danthine, 1981). The purpose of introducing the concept of cross-hedging in the analysis is that the dynamic procedures developed in this thesis are likely to be beneficial in cross-hedging scenarios - in such situations the relationship between spot and futures prices is likely to be less than perfect. This section discusses the issues involved when hedging must be accomplished by adopting a futures position in an asset that is not exactly identical to the underlying. In such instances, effective hedges are likely to be time-varying in nature.

As mentioned in the previous section, a major problem associated with hedging is encountered in the presence of lightly traded, relatively underdeveloped or non-existent futures markets. In situations where basis risk possesses time-varying characteristics, dynamic techniques are likely to offer greater risk-reduction alterna-
tives to static procedures. The potential efficiency of a cross-hedge can be measured through correlation analysis. Futures markets provide the opportunity to yield very good hedges when the financial instrument in the spot market has a futures contract that exactly mirrors its characteristics - the correlation between spot and futures prices is likely to be high. However, in empirical situations hedging must often be accomplished via the use of futures contracts that deliver a different asset, liability or derivative to the underlying spot market instrument.\textsuperscript{26} In such cases, hedging involves selection of an appropriate futures contract that displays similar price fluctuations with the spot market instrument, and the subsequent determination of the proportion of the investment to be hedged using futures contracts or options.

The problem involved in cross-hedging is that basis risk may be larger since the correlation between spot and futures prices may not be as strong. For example, the credit rating of the hedged spot instrument may not be as high as the credit worthiness of the instrument reflected in the futures contract, or the liquidity of both spot and futures instruments may not be similar, or the spot instrument may have a limited supply. As a result, spot market prices behave differently to the futures price for any chosen contract. The futures price tracks changes in the spot market instrument to which it actually relates. There is no reason to expect the futures price to converge to the spot price, even on the maturity date of the futures contract. Cross-hedging is therefore considered riskier than direct hedging because of this additional dimension of basis risk. The issue of structuring effective cross-

\textsuperscript{26}For example, corporate bonds may be held and since there may not exist a futures market for these bonds, the only alternative is to hedge using treasury bond futures.
hedges is discussed in this section.

2.5.1 Effective Cross-hedging

Having outlined the concept of cross-hedging, the focus turns to the structuring of effective cross-hedges. There are several issues that impact on the ability to construct effective cross-hedges that should be examined before effectuating such a hedge. The relationship of this section to the overall thesis lies in the provision of guidelines to the hedger that facilitate an effective cross-hedge strategy.

There are several issues that may influence the effectiveness of a cross-hedge. They include, but are not limited to:

1. the degree to which spot and futures prices are positively correlated. The problem of finding a substitute asset or commodity whose price movements follow as closely as possible to those of the asset or commodity of interest is vitally important and is essentially a correlation problem. If the futures market in a certain commodity is thin (or the market is not actively traded), the hedger cannot cover themselves well (by effectively replacing price risk with the more predictable basis risk), leading to basis risk being unacceptably high and the hedge not being effective.\(^{27}\) The market itself might have large premiums (spreads) or may be extremely volatile, even in futures markets. In other words, the relationship between spot and futures prices may be very volatile over time.

\(^{27}\)The month in which the hedger would like to buy the asset or commodity may not be available. Even where the hedger can obtain this chosen month, they may not be able to close out the position.
2. the stability of the cross-hedge ratios over time (Aggarwal and DeMaskey, 1997; DeMaskey, 1997; Broll, Mallick and Wong, 2001). A reasonable expectation in situations of direct hedging is that the correlation between spot and futures prices is very stable because of the non-arbitrage pricing relationship. However, in cross-hedging, no such relationship exists. The correlation between the spot asset and the hedging instrument may exhibit substantial time-variation. Instability of the hedge ratios usually leads to reduced hedging effectiveness. Such instability raises concerns over the usefulness and reliability of the cross-hedge to reduce price risk.

The technique of cross-hedging has been found to be of practical benefit in hedging various instruments such as mortgage-backed securities and treasury bills (Koutmos, Kroner and Pericli, 1998), currency futures (Broll, Wong and Zilcha, 1999) and commodities (Dahlgran, 2000). The next task is to discuss foreign exchange cross-hedging as this type of hedging is common in the Asian region, where the exposure to adverse movements in many currencies are hedged via currency futures contracts denominated in the Japanese Yen.

### 2.5.2 Foreign Exchange Cross-Hedging

The dissertation deals with several examples involving foreign exchange rates. It therefore appears logical to provide an overview of foreign exchange hedging and the difficulties involved when hedging using lightly traded or thin foreign exchange futures markets. In such circumstances, effective hedging may be accomplished by
selecting a currency that is highly correlated with the underlying currency of interest. It should be expected that such cross-hedging results in lower effectiveness than otherwise achievable. This section summarises foreign exchange rate risk (currency risk) and the need for hedging in such markets.

Risk management in foreign exchange markets is necessary since spot exchange rates between currencies may be very volatile, even in the short-run. Corporations may find that between the time of agreeing on payment for an overseas transaction and the time of actually paying the debt, there may have been a significant adverse movement in exchange rates. Therefore, corporations who expect to make (receive) payments in terms of a foreign currency at a future date face the risk they have to pay more (receive less) in terms of the domestic currency than anticipated. In such situations the major issue for corporations is the effective and efficient hedging of these foreign exchange exposures through currency futures contracts.

In certain instances, currency risk may not be easily hedged in international currency markets, forcing the hedger to implement a cross-hedge. Currency cross-hedging is a method adopted with the objective of minimising exchange rate risk when the expected cash flows are denominated in a minor currency (DeMaskey, 1997). Aggarwal and Mougoue (1998) have found evidence of a nascent Yen bloc in the Asian-Pacific region with many currencies of the Asian newly industrialised

---

28Hopper (1997) concluded that exchange rates do not seem to be affected by economic fundamentals such as money supplies and central bank policies, but rather are impacted by market sentiment in the short-run. As a result, the most accurate forecast of the exchange rate in the short-term is the current exchange rate.
countries (NICs) being cointegrated with the Japanese Yen, implying that currency risks in these Asian NICs may be hedged using instruments denominated in the Japanese Yen or other major currencies. Aggarwal and DeMaskey (1997) concluded that cross-hedging risk management strategies involving derivatives denominated in more liquid currencies may be beneficial in Asian emerging markets by offering improved risk-return combinations over unhedged portfolios.\footnote{Currency hedging is unlikely to be beneficial when purchasing power parity and interest rate parity conditions hold (Aggarwal and DeMaskey, 1997).}

Hedging foreign exchange risk in developing (emerging) markets may be difficult, if not impossible, as derivative markets in such currencies are either non-existent or relatively underdeveloped, or exchange controls and other government regulations restrict the activities of the foreign exchange market in these currencies (DeMaskey, 1997). While derivative contracts may be available in these minor currencies, such contracts tend to be thin, less liquid at increasing contract maturity and more volatile than those of developed markets. In such instances, exposure to foreign exchange risk may be cross-hedged using a derivative contract in a liquid currency highly correlated with the underlying exchange rate. However, this may lead to further hedge ratio instability since the underlying structural factors of the cross-hedge currency involved may not be well-developed (Grammatikos and Saunders, 1983), leading to the correlation between the currency and the cross-hedging instrument being weak and less stable.

This section highlighted the importance of selecting a cross-hedging instrument whose prices are highly correlated with those of the instrument being hedged. The
hedge ratios must be constantly monitored going forward and updated if they are found to lose their effectiveness over time. The monitoring of the hedge ratios is more important in the cross-hedging situation than in the direct hedging scenario due to the higher possibility the hedge becomes less effective over time. Furthermore, the availability of various alternative hedging instruments may reduce the exposure to adverse price movements in the underlying instrument more than the current hedge.

### 2.6 Conclusion

This chapter discussed the mechanics behind the concept of hedging, an understanding of which provides the foundation for examining the potential role of conditional volatility hedging models. The main purpose of hedging is to replace the sometimes volatile spot risk with the less volatile basis risk. While the time series behaviour of spot and futures returns may display substantial volatilities, the pricing relationship between the two usually results in the time series of the basis being much more stable. If the basis spread is adequately modelled and basis risk reduced to an acceptable level, the hedge is likely to be effective. Therefore, the modelling of the basis is critical in maximising hedging effectiveness.

The basis is assumed to be small and stable over time - an assumption that traditionally underlies the models of time series behaviour used in hedge ratio calculations. Violation of this assumption may impair the calculation and maintenance of the hedge as the spread volatility between spot and futures prices changes significantly over time. Therefore, there would appear to be a real issue involving the
possibility of time-dependent basis volatility - this issue is ignored when applying minimum-variance or other constant hedges. Omission of time-varying volatility may lead to hedge ratios that are less effective than otherwise possible. The technique of cross-hedging was outlined since basis risk is likely to be substantial in such a situation.
Chapter 3

Conditional Variance

Specifications

3.1 Introduction

The thesis investigates different conditional variance specifications that may adequately capture dynamic characteristics present in financial time series. While the context of the analysis is the spread between spot and futures prices, the research applies to any spread between two price series. The proposition underlying the research is that short-term information alone may not adequately capture the dynamic characteristics of spot and derivative price series over time. Cointegration offers a means of incorporating long-run information into current volatility specifications such as GARCH models. This chapter reviews the relevant distributional characteristics of financial asset returns as a prelude to the discussion of GARCH models in Chapter 4. The distributional characteristics of asset returns include skewness,
leptokurtosis and volatility clustering and are all captured to varying degrees by the GARCH family of models.

This chapter is organised as follows: Section 3.2 discusses the key distributional characteristics exhibited by financial asset returns. The section also outlines several methods for testing whether the variance of the returns is in actual fact time-varying. Section 3.3 summarises one such conditional variance framework: the autoregressive conditional heteroscedastic (ARCH) specification.

While ARCH models are effective in capturing the volatility in the conditional second moments of asset returns, they have limitations. A more generalised framework, known as the GARCH specification, provides an even more comprehensive model for capturing the conditional variance. GARCH models permit investigation of both the transitory and permanent impact of shocks to the system on the variance. The GARCH methodology in both a univariate and a multivariate framework is outlined in Section 3.4, including specification of various forms of the multivariate GARCH model.

Section 3.5 extends the GARCH formulation by incorporating cointegration information into the specification of the conditional second moments of the returns. Cointegration may assist in superior modelling of the conditional variance by including a term that allows for any transitory disequilibrium between spot and futures prices. The section also outlines the procedures commonly invoked in the estimation of parameters from the GARCH model. Section 3.6 summarises the issues discussed in this chapter and identifies the analytical steps that follow in the research.
3.2 Distributional Characteristics of Returns

The prescriptions of modern financial risk management hinge critically on the associated characterisations of the distribution of spot and futures returns (Andersen, Bollerslev, Diebold and Labys, 2001). A summary of these characteristics is provided before the consideration of potential conditional variance models that capture the distributional characteristics of the returns. Since the original work of Mandelbrot (1963) and Fama (1965), it has been accepted the time series features commonly exhibited by financial asset returns - skewness, leptokurtosis, volatility clustering - violate normality assumptions, rendering standard Gaussian tests sub-optimal.\(^{30}\) These characteristics are also present in the asset prices simulated in several examples in the thesis.

The most common characteristics of financial asset returns are discussed along with the tests that may be carried out to determine whether financial returns possess such characteristics. GARCH and GARCH-X specifications are considered in this analysis in modelling the volatility of, and between, spot and futures returns, aiming to capture the distributional characteristics. The purpose of this part of the analysis is to provide the context and knowledge necessary for understanding subsequent development of the GARCH and GARCH-X models. The first task involves a discussion of the notions of skewness and kurtosis and their statistical measures. Volatility clustering is also defined, followed by a discussion of the logic behind the

\(^{30}\)While the unconditional distributions of high-frequency asset returns have been well-known to exhibit substantial skewness and/or excess kurtosis when compared to the normal distribution, the conditional distributions of these asset returns are normal.
3.2.1 Skewness

A characteristic often observed in financial time series is skewness. Skewness is defined as the third centralised moment of a distribution (the average cubic departure from the mean) and measures the direction and degree of asymmetry of the distribution. A value of zero indicates a symmetric distribution. A distribution with a positive (negative) value for its skewness displays a longer tail on the right (left) side of the distribution, and its mean is typically greater than (less than) its median, which in turn is greater than (less than) its mode. Since the mean exceeds (does not exceed) the median, most of the returns are below (above) the mean, but they are of smaller (larger) magnitude than the smaller (larger) number of outcomes that are above (below) the mean.

Mathematically, the coefficient of skewness is defined as

$$\beta_1 = \frac{E[(X - \mu)^3]}{\sigma^3},$$

where the numerator is the third centralised moment, $\mu = E[X]$, $\mu$ is the mean and $E[X]$ denotes the expected value of the random variable $X$.

3.2.2 Kurtosis

A second characteristic often observed in most financial time series is kurtosis. Not all ARCH/GARCH type models do equally well in capturing kurtosis (Brooks, Burke and Persand, 2001). Kurtosis is defined as the fourth centralised moment of a
distribution (the average quadric departure from the mean) and measures the degree of peakedness and heaviness of the tails of the distribution. A distribution with wide tails and a tall narrow peak possesses excess kurtosis and is called leptokurtic. The kurtosis value of such a distribution is significantly greater than three. Compared with a normal distribution, a larger fraction of the returns are at the extremes rather than slightly above or below the mean of the distribution. On the other hand, a distribution that possesses thin tails and a relatively flat, wide centre has a kurtosis value of less than three. Relative to the normal distribution, a larger fraction of the returns are clustered around the mean.

Mathematically, the coefficient of kurtosis is defined as

$$\beta_2 = \frac{E[(X - \mu)^4]}{\sigma^4},$$

where the numerator is the fourth centralised moment.

### 3.2.3 Volatility Clustering

A third characteristic exhibited by most financial time series is volatility clustering. Where volatility clustering characterises a time series, the ARCH/GARCH family is most often appropriate (Alexander, 2001). Volatility clustering is the tendency for large (small) values of past squared returns to give rise to large (small) current values, with more or less smooth transition from higher (lower) to lower (higher) volatility. In other words, the current level of volatility tends to be positively correlated with its level during the immediately proceeding periods. In the application of traditional econometric models, volatility clustering causes problems since the
clustering usually transmits to the variability of the returns, resulting in its heteroscedasticity. Therefore, volatility clustering is an integral part in characterising the conditional density function of asset returns and should be considered in the construction of conditional distributions. If volatility clustering is the primary factor in explaining price changes, conditional volatility models are superior to any independent models. Substantial dependence would not exist in the data after removing the ARCH/GARCH type effects.

ARCH/GARCH models are referred to as dependent models since they do not have any independent random variables. In their original form, the ARCH and GARCH models assume a normal distribution, with a conditional variance that changes over time. For the ARCH model, the conditional variance varies over time and is a function of past squared deviations from the mean. The conditional variance of the GARCH process changes over time as a function of past squared deviations from the mean and previous variances. In contrast, for the GARCH-X model the conditional variance changes over time as a function of not only past squared deviations from the mean and previous variances, but also as a function of the squared error correction term. Lee (1994) proposes that examining the behaviour of the variances over time as a function of short-run deviations is reasonable when one expects increased volatility, due to shocks to the system which propagate on the first and second moments.

This chapter shows the GARCH family of models are superior to ARCH mod-

---

31 An example of an independent model is a mixture of normals, since it assumes the observations are independent random variables.
els in their capability of dealing with departures from normality that dominate the behaviour of many time series. As a result of this dominance, the thesis adopts the GARCH framework in a bivariate setting as the platform in modelling the conditional variance and covariance of financial assets and, in turn, constructing the dynamic hedge ratios. The ARCH specification is firstly introduced and then generalised to obtain the GARCH model. The ARCH and GARCH models are defined in Section 3.3 and Section 3.4 respectively. The next task in the development of the model is to test for possible departures from normality – this step is explained in the next section.

3.2.4 Testing for Departures from Normality

This section provides several tests to determine whether, and the extent to which, financial asset returns exhibit non-normal characteristics. Such testing is necessary to validate the use of conditional variance type models. Conditional variance models should be invoked if the returns series deviate significantly from normality. Parametric tests of a model’s parameters usually require an assumption of normality in the distribution of residuals.

In order to conduct single or joint hypothesis tests of the model parameters, the disturbances \( \{e_t\} \) from any postulated linear regression of the form

\[
y_t = \beta_0 + \beta_1 x_{1t} + \ldots + \beta_n x_{nt} + e_t
\]

should follow a normal distribution with mean 0 and (constant) variance \( \sigma^2 \). Various tests may be applied to the residual sequence \( \{e_t\} \) to examine any possible departures
from normality. Among the most common tests are the Jarque-Bera normality test of Jarque and Bera (1980), the Ljung-Box autocorrelation test (Ljung and Box, 1978) and the Lagrange-Multiplier test for ARCH effects (White, 1980). The three tests are now explained prior to their use for calculating exploratory statistics in the examples that follow in later chapters.

The Jarque-Bera Normality Test

Jarque and Bera (1980) noted that since the normal distribution is fully described by the first two centralised moments - the mean and the variance - a test may be devised to evaluate the null hypothesis the sequence \( \{e_t\} \) has a normal distribution with unspecified mean and variance. For a normal distribution, both the mean and variance should be constant. The alternative is the process \( \{e_t\} \) does not have a normal distribution. Denote the sample size as \( N \), the skewness as \( b_1 \) and the excess kurtosis as \( b_2 \), both of which may be estimated using \( \{e_t\} \). The Jarque-Bera test determines whether the sample skewness and kurtosis are significantly different from their expected values, as measured by a \( \chi^2 \) (chi-squared) statistic. The Jarque-Bera procedure jointly tests \( b_1 = 0 \) and \( b_2 = 0 \).

Under the null hypothesis of normality of the residuals, the test statistic is asymptotically distributed as a \( \chi^2 \) random variable with two degrees of freedom. The Jarque-Bera statistic is given by

\[
JB = \left( \frac{N}{6} \right) b_1^2 + \left( \frac{N}{24} \right) b_2^2
\]
The Ljung-Box Test

Unlike the Jarque-Bera test, the Ljung-Box ($Q$) test examines the null hypothesis that all the correlation coefficients ($\rho_k$) in Equation (3.1) are simultaneously equal to zero (white noise).\textsuperscript{32} If all correlation coefficients are equal to zero, the residuals \{\(e_t\}\} should form a white noise process with constant mean, constant variance and uncorrelated sequential observations. If the alternate hypothesis is true, at least one of the $\rho_k$ is not equal to zero - the process is not white noise. The $Q$ test is an improved modification of the Box-Pierce test developed by Box and Pierce (1970). The test examines the first $m$ lags in the autocorrelation function of the residuals \{\(e_t\}\} and may be applied to both the absolute values and the squares of the residuals to examine whether these processes are white noise or linearly dependent.

Under the null hypothesis of white noise of the residuals, the test statistic is asymptotically distributed as a $\chi^2_{m-k}$ distribution and is defined by

$$Q = N(N + 2)\sum_{k=1}^{m} \frac{1}{N-k} \rho_k^2$$ (3.1)

The Lagrange-Multiplier Test

The Lagrange-Multiplier test evaluates the null hypothesis of no ARCH effects. A test for ARCH of order $q$ involves regressing the squared residuals, $e_t^2$, on a constant and the $q$ lags of $e_t^2$. Mathematically,

$$e_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \ldots + \alpha_q e_{t-q}^2 + \eta$$ (3.2)

\textsuperscript{32}A white noise process has constant mean, constant variance and each observation is uncorrelated with all other values in the sequence.
The coefficient of multiple correlation, $R^2$, is obtained from the regression in Equation (3.2). The test statistic is defined as $NR^2$ (the number of observations multiplied by the coefficient of multiple correlation). Under the null hypothesis of no ARCH effects, $\alpha_1 = \ldots = \alpha_q = 0$. The resulting joint effects should be slight or non-existent. In other words, the variance is constant. If the null is rejected, the squared residual term can be explained by its own lag and the variance is therefore non-constant. The test statistic is asymptotically distributed as a $\chi^2$ random variable with $m$ degrees of freedom. The alternative hypothesis is that at least one of the coefficients $\alpha_1, \ldots, \alpha_q$ is not equal to zero.

In this thesis, the Jarque-Bera, Ljung-Box and Lagrange-Multiplier tests all precede the application of conditional variance models. Significance in one or more of these tests implies lack of residual normality and the potential for applying time-varying procedures. Standard Gaussian models are unlikely to be of benefit when the departures from normality are significant. If the returns exhibit significant non-normal characteristics, the data is a candidate for ARCH/GARCH type modelling. The conditional variance specifications examined in the thesis are the bivariate GARCH and bivariate GARCH-X models. Before discussing these models, the ARCH specification is introduced in the following section.

### 3.3 ARCH Models

The framework implemented in the thesis to model the stylised features of asset returns are GARCH-type models. Since GARCH models are derived from ARCH
models these latter models are explained first. The generalisation of the ARCH framework (the GARCH formulation) follows in the next section. The aim of the review is to identify suitable time-varying models that allow for dynamic distributional properties of asset returns. Time-varying variances and covariances may be constructed and used to calculate conditional hedge ratios that are likely to be superior to static alternatives. This section mathematically defines the ARCH model, discusses its characteristics and demonstrates why variations of such a specification are important in modelling asset returns, providing the logic for the practical application of similar models in later chapters.

In a seminal paper, Engle (1982) argued that economic agents respond not just to the mean, but also to higher order moments of random variables. Therefore, the mean and variance of the returns of financial assets may be significant for portfolio decisions. Engle (1982) proposed the ARCH framework to model the time-varying conditional variance. ARCH models are superior to other techniques in the finance context since these models are deterministic volatility specifications, able to capture the heteroscedasticity of the returns of financial time series. Derivations of such models are frequently invoked in the estimation of conditional variances and covariances of asset returns.

The ARCH framework allows two distinct specifications: a conventional regression specification for the mean function, with a variance specification that is permitted to change over the sample period. The ARCH model expresses the current conditional variance of a time series of total returns as a weighted average of its past squared asset returns. Greater weight is placed on more recent rather than on more
distant squared returns in an attempt to capture the volatility clustering present in
the returns. Bollerslev, Engle and Nelson (1994, pg. 1) state that “parallel to the
success of standard linear time series models, arising from the use of the conditional
versus the unconditional mean, the key insight offered by the ARCH model lies in
its distinction between the conditional and unconditional second-order moments”.
The term conditional implies explicit dependence on a past sequence of observa-
tions. The term unconditional is more concerned with long-term behaviour of a
time series and assumes no explicit knowledge of the past.

3.3.1 The ARCH(q) Model

ARCH models recognise the presence of successive periods of relative volatility and
stability. The error variance, conditional on past information, evolves over time as a
function of past errors. ARCH models can be implemented to predict changes in the
conditional variance of asset returns. In the univariate ARCH model, the variance
of a zero-mean random variable at time \( t \), conditional on information available at
that time, is assumed to be a function of past values of that random variable and
some unknown set of parameters. Assume that

\[
e_t = z_t \sigma_t, \tag{3.3}
\]

where \( z_t \) is a standardised, independent and identically distributed random variable
such that

\[
E(z_t) = 0 \quad \text{and} \quad Var(z_t) = 1 \tag{3.4}
\]
The formulation in Equations (3.3) and (3.4) implies the conditional and unconditional mean of \( \{ e_t \} \) is zero. Although successive innovations are uncorrelated, they are not independent. Denoting the information set by \( F_{t-1} \), Engle (1982) expressed a parameterisation for the conditional variance of \( \{ e_t \} \) as a linear function of past squared values of the model, that is,

\[
\begin{align*}
\text{Var}(e_t | F_{t-1}) &= \sigma_t^2 \\
&= \alpha_0 + \sum_{j=1}^{q} \alpha_j e_{t-j}^2,
\end{align*}
\]

where \( \alpha_0 \) is a constant, \( q \) is the order (number of linear regression components) of the model and \( \alpha_j > 0 \) (for \( j = 1, \ldots, q \)) are the coefficients of the linear regression parameters (and ensure the model is well-defined and the conditional variance non-negative).

In Equation (3.5) the \( \{ e_t \} \) are independently and identically distributed with means 0 and variances \( \alpha_1, \ldots, \alpha_q \) respectively. The specification is known as the ARCH\((q)\) model. The process is covariance stationary if and only if the sum of the positive autoregressive parameters is less than one. The resultant unconditional variance is equal to

\[
\begin{align*}
\text{Var}(e_t | F_{t-1}) &= \text{E}(e_t^2) = \sigma_t^2 \\
&= \frac{\alpha_0}{1 - \alpha_1 - \ldots - \alpha_q}
\end{align*}
\]

Note that ARCH models are consistent with various forms of efficient market theory, which states that asset returns observed in the past cannot improve the forecast of asset returns in the future.\footnote{Since the innovations \( \{ e_t \} \) are serially uncorrelated, ARCH modelling does not violate efficient market theory.} In practical situations, a major problem is
encountered in fitting ARCH($q$) models to financial returns data, in that the order $q$ has to be relatively large in order to obtain a good-fitting model. If $q$ is large, this requires the estimation of a large number of parameters. The $\{e_t\}$ in the original formulation of the ARCH model is limited to a univariate process. For an extension of the univariate ARCH representation to a multivariate ARCH specification, see Engle and Kraft (1983).

This section reviewed ARCH models. While the ARCH specification does allow the error variance to evolve over time as a function of past errors, it has the limitation of not incorporating past conditional variances in the expression for the current conditional variance. The second alternative foundation is the GARCH family of models. GARCH models may dominate ARCH models in their ability to better incorporate the departures from normality frequently observed in financial time series.

### 3.4 GARCH Models

The objective of this part of the analysis is to review alternative time series models that may be applied in a hedging framework to counter the risk in an existing spot position. This section provides an overview of the GARCH framework and lists several characteristics of such a formulation. As emphasised by Bollerslev, Engle and Nelson (1994), many financial time series have a number of characteristics in common, for example, although return series usually show little or no autocorrelation,
serial independence between the squared values of the series is likely to be rejected since there may exist nonlinear relationships between subsequent observations.

To reduce the computational burden of the ARCH model, Bollerslev (1986) generalised the ARCH framework by providing a parsimonious, more flexible and comprehensive representation of the lag structure, known as the generalised autoregressive conditional heteroscedastic (GARCH) model. The conditional variance in a GARCH model is a function of a long-term average (the constant), the forecasted variance from the last period (the GARCH term) and information about volatility observed in the previous period (the ARCH term). GARCH models have been implemented successfully in modelling time-varying volatilities and correlations by not only incorporating the heteroscedasticity observed in economic and financial data but also capturing any skewness and leptokurtosis present in such data (Bollerslev, Chou and Kroner, 1992). While similar to ARCH models, the GARCH family has advantages for capturing the distributional characteristics of basis spreads.

Unlike ARCH models, the GARCH($p, q$) model allows for the decomposition of the conditional covariance into two components (Darbar and Deb, 1997). The first, permanent (long-run) component, is the historical (unconditional) mean covariance between these returns. A fluctuation of the conditional covariance around this permanent component is the second, transitory (short-run) component of the covariance at any point in time. The GARCH($p, q$) model may be written as

\[
Var(e_t|\mathcal{F}_{t-1}) = \sigma_t^2 \\
= \alpha_0 + \sum_{j=1}^{q} \alpha_j e_{t-j}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2,
\]

(3.6)
where $\alpha_0$ is a constant, and $\alpha_j > 0$ for $j = 1, ..., q$ and $\beta_i > 0$ for $i = 1, ..., p$ are coefficients of the $e_{t-j}^2$ and $\sigma_{t-i}^2$ terms.

Volatility reversion implies that past volatility shocks have a transitory effect on future volatility, in that their impact diminishes over time. A stochastic process $x_t$ is called weak or covariance stationary if $E(x_t)$ is constant, $Var(x_t)$ is constant and, for any $t$ and $f \geq 1$, $Cov(x_t, x_{t+f})$ depends only on $f$ and not on $t$.\textsuperscript{34} If the time series is covariance stationary, any movements away from the permanent level of covariance are transitory - the covariance tends to return to the permanent level. If the conditional variance contains a substantial transitory component, the impact of a shock fades slowly and the time path of the hedge ratio requires constant revision. Due to the nonlinearities caused by GARCH effects, the hedge ratio changes as the hedging horizon changes.

A unit root in the volatility process suggests the conditional variance tracking the time path of the variance may exhibit both a trend and a cyclical component. A unit root in the volatility process has important implications for the time-dependant hedge ratio. If the variance undergoes a shock, its time path shifts and the time path of the hedge should be adjusted. The distinction between permanent and transitory components of the covariance is necessary since effective diversification strategies may be based on the level and dynamics of the conditional covariance, that are likely to be altered by short-run volatilities, resulting in the basis being volatile. If

\textsuperscript{34}The weaker form of stationarity requires only the mean and variance to be constant over time and the covariance between two components is a function only of the lag between them. A covariance stationary process of uncorrelated components is called white noise.
transitory changes in covariance away from the permanent level are large and where
the covariance does not return to the permanent level immediately - that is, if the
covariance returns to the permanent level after several trading days - adjustments
to a portfolio based on these temporary covariances may be beneficial.

The conditional variance in Equation (3.6) is well-defined when all the coeffi-
cients in the corresponding infinite order ARCH model are positive. The necessary
and sufficient conditions for such non-negativity are outlined in Nelson and Cao
(1992). If the $\beta_i$ in Equation (3.6) equal zero for all $p$ lags, the conditional variance
of Bollerslev’s GARCH($p, q$) process reduces to Engle’s (1982) ARCH($q$) process.
Furthermore, $\sigma_t^2$ is finite and $\{e_t\}$ is a covariance stationary process when

$$\sum_{i=1}^{p} \beta_i + \sum_{j=1}^{q} \alpha_j < 1$$

The above is a sufficient but not necessary condition for strict stationarity (Boller-
slev, Engle and Nelson, 1994).

### 3.4.1 The GARCH(1,1) Model

Bollerslev, Chou and Kroner (1992) find in most applications $p = q = 1$ suffices.
The GARCH(1,1) model is given by

$$\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \quad (3.7)$$

where $\alpha_0$ is a constant, $\alpha_1$ indicates the contribution to conditional variance of the
most recent news, and the parameter $\beta_1$ corresponds to the moving average part in
the conditional variance (that is, the recent level of volatility). The variance forecast,
\( \sigma_t^2 \), consists of a constant plus a weighted average of last period’s forecast, \( \sigma_{t-1}^2 \), and last period’s squared disturbance, \( \epsilon_{t-1}^2 \). The conditions ensuring non-negativity and stationarity amount to ensuring that both \( \alpha_1 \) and \( \beta_1 \) are non-negative (Nelson and Cao, 1992). Resultant shocks to volatility fade out at the constant rate of \( (\alpha + \beta) \).\(^{35}\)

In such a situation, the GARCH(1,1) model corresponds exactly to an infinite order linear ARCH model with geometrically declining parameters.

The GARCH(1,1) model, due to its employment of the square of the past residual, is well suited to capture the phenomenon of volatility clustering. However, Baillie and Bollerslev (1989) and Hsieh (1989) note that significant leptokurtosis still exists even after fitting the GARCH(1,1) model, though such a model adequately captures volatility clustering. Note that when \( 0 < \alpha_1 + \beta_1 < 1 \), \( \epsilon_t \) has an unconditional stationary distribution that is non-normal with variance equal to

\[
\sigma^2 = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}
\]

According to Poterba and Summers (1986), the impact of volatility can only be significant if shocks to volatility persist over a long time. The market does not make an adjustment of the future discount rate if shocks to volatility are not permanent in nature. Expected returns are not affected by the volatility movement if shocks to volatility are transitory. If \( \alpha + \beta = 1 \), the variance would contain a unit root. Such a model is termed an Integrated GARCH or IGARCH model. In other words, when \( \alpha + \beta = 1 \), the current shock persists indefinitely in conditioning the future variance.

\(^{35}\)After \( n \) periods, the proportion of the initial shock that remains is \( (\alpha + \beta)^n \).
In a GARCH(1,1) model, the \( \{e_t\} \) is covariance stationary if the sum of \( \alpha \) and \( \beta \) is significantly less than unity. As the sum of \( \alpha \) and \( \beta \) approaches unity, the shock persists for a relatively long time. Since \( \alpha + \beta \) represents the change in the response function of shocks to volatility per period, a value greater than unity implies the response function of volatility increases with time, and a value less than unity implies that shocks decay over time. The closer the persistence measure is to unity, the slower the decay rate.

### 3.4.2 Multivariate GARCH Models

The success of GARCH models in fitting univariate time series has motivated researchers to extend these models to the multivariate situation (Wahab, 1995). Simplicity is the main advantage of the univariate GARCH model as it utilises only the information in a market’s own history. Generalising a univariate GARCH model to a multivariate GARCH (MGARCH) specification is achieved by allowing the entire variance-covariance matrix to change over time, rather than just the variance. Therefore, the elements of the covariance matrix are a linear function of lagged squared errors, lagged cross-products of the errors, lagged variances and lagged covariances.

The residuals in an MGARCH specification are assumed to follow a conditional multivariate normal distribution with mean 0 and variance-covariance matrix \( H_t \). Conrad, Gultekin and Kaul (1991) concluded that multivariate models lead to more precise estimates of the parameters because they utilise the information in the entire variance-covariance matrix of the residuals and allow the variance-covariance
to depend on the information set in an ARMA manner. Furthermore, the generated regressor problem associated with univariate models is avoided in multivariate models because the latter estimate all parameters jointly (Pagan, 1984).

The conditional variance equations in MGARCH models are able to capture volatility clustering and provide a means for exploring the transmission of volatility shocks from one market to another, known as volatility spillovers. MGARCH models may be implemented to identify both the source and magnitude of such volatility spillovers. These models are potentially useful in regards to the parameterisation of conditional cross-moments. Examples of MGARCH models are the VEC representation of Bollerslev, Engle and Wooldridge (1988) and the BEKK specification of Baba, Engle, Kraft and Kroner, as proposed by Engle and Kroner (1995). Both of these formulations are outlined, and the BEKK model implemented, in Chapter 4.

The VEC Parameterisation

This section introduces the multivariate linear GARCH($p, q$) VEC parameterisation of Bollerslev, Engle and Wooldridge (1988). The VEC framework assumes that $H_t$ is specified as a linear function of the lagged cross-squared errors and lagged values of $H_t$.\textsuperscript{36} Assume that

$$e_t | \mathcal{F}_{t-1} \sim (0, H_t),$$

\textsuperscript{36}For the VEC parameterisation to yield estimates that guarantee a positive semi-definite $H_t$ matrix, the restriction that all estimated elements are greater than or equal to zero would have to be imposed.
where $e_t$ represents the innovation in the conditional mean equation, and where VECH(,) denotes the vector-half operator that stacks the lower triangular elements of an $N \times N$ matrix into an $\left\lceil \frac{N(N+1)}{2} \right\rceil \times 1$ vector. Denoting

$$e_{t-1} = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix},$$

the bivariate VEC GARCH(1,1) model may be expressed as

$$VECH(H_t) = C_0 + AVECH(e_{t-1}e_{t-1}') + BVECH(H_{t-1}),$$

where

$$VECH(H_t) = \begin{pmatrix} H_{11,t} \\ H_{12,t} \\ H_{22,t} \end{pmatrix},$$

$$VECH(e_{t-1}e_{t-1}') = \begin{pmatrix} u_{11,t-1}^2 + u_{12,t-1}^2 \\ u_{21,t-1}^2 + u_{22,t-1}^2 \\ u_{11,t-1}u_{21,t-1} + u_{12,t-1}u_{22,t-1} \end{pmatrix},$$

$$C_0 = \begin{pmatrix} c_{11} \\ c_{12} \\ c_{22} \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}.$$

The conditional variances and conditional covariances depend on the lagged values of the conditional variances of asset returns in the series and the conditional covariances between them, as well as the lagged squared errors and the error cross-products. While the VEC representation is very general, such a specification is
not parsimonious, it involves a large number (21) of parameters for all values of $e_t$, subject to the requirement that $H_t$ is positive-definite.\(^{37}\)

Further restrictions on the parameterisation may be desirable in practice. A restriction on the matrices $A$ and $B$ to be diagonal results in the model proposed by Bollerslev, Engle and Wooldridge (1988). In this model, each element in the variance-covariance matrix is postulated to follow a GARCH-type equation with the lagged variance-covariance term and the product of the corresponding lagged residuals as the variables in the conditional covariance equation. The diagonal form of the VEC specification may be expressed as

$$
V ECH(H_t) = C_0 + A V ECH(e_{t-1} e_{t-1}^t) + B V ECH(H_{t-1}),
$$

where

$$
V ECH(H_t) = \begin{pmatrix}
H_{11,t} \\
H_{12,t} \\
H_{22,t}
\end{pmatrix},
$$

$$
V ECH(e_{t-1} e_{t-1}^t) = \begin{pmatrix}
u_{11,t-1}^2 + u_{12,t-1}^2 \\
u_{21,t-1}^2 + u_{22,t-1}^2 \\
u_{11,t-1} u_{21,t-1} + u_{12,t-1} u_{22,t-1}
\end{pmatrix},
$$

\(^{37}\)The number of parameters in the variance-covariance system increases to 78 in the case of 3 equations and 210 when considering 4 equations. Therefore, the general model becomes very difficult to estimate and analyse as the number of parameters increases dramatically.
In such a specification, there are nine parameters to be estimated in the case of two equations and 18 parameters in the case of three equations. Where the matrices $A$ and $B$ are further restricted to be diagonal, the conditional variances are a function of their own lagged values and their own lagged residual terms, while the conditional covariance is a function of lagged covariances and lagged cross-products of the $\{e_t\}$. The diagonal restriction is often enforced in practice as it results in a more parsimonious representation of the conditional variance (Bollerslev, Engle and Nelson, 1994).

The BEKK Parameterisation

A major disadvantage of the VEC representation is that, unless the model is constrained in some manner, there is no guarantee of a positive semi-definite $H_t$ matrix. Due to this limitation, an alternative MGARCH parameterisation (known as the BEKK parameterisation of Engle and Kroner, 1995) is adopted in the dissertation. The BEKK model overcomes the difficulties associated with the VEC framework. Such a representation is expressed as

$$H_t = C'C + A'e_{t-1}e_{t-1}'A + B'H_{t-1}B,$$

(3.8)
where

\[
H = \begin{pmatrix} H_{11} & H_{21} \\ H_{21} & H_{22} \end{pmatrix}, \quad C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix},
\]

\[
A = \begin{pmatrix} A_{11} & A_{21} \\ A_{21} & A_{22} \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & B_{21} \\ B_{21} & B_{22} \end{pmatrix},
\]

\[
e_{t-1}e'_{t-1} = \begin{pmatrix} u_{11,t-1}^2 + u_{12,t-1}^2 & u_{11,t-1}u_{21,t-1} + u_{12,t-1}u_{22,t-1} \\ u_{21,t-1}u_{11,t-1} + u_{22,t-1}u_{12,t-1} & u_{21,t-1}^2 + u_{22,t-1}^2 \end{pmatrix}.
\]

One important feature of the BEKK specification is that it is sufficiently general to allow the conditional variances and covariances of the two markets to influence each other. The structure of Equation (3.8) shows the BEKK parameterisation retains high generality by allowing for a wide range of possible dynamic interaction between the conditional variances of the returns and the conditional covariances of the returns. Secondly, the BEKK model also imposes the condition of positive-definiteness on the estimated conditional variance-covariance matrix. Weak restrictions on \( A \) and \( B \) guarantee that \( H_t \) is always positive-definite. Thirdly, the BEKK model allows for dynamic correlation between the returns.

A fourth attribute/characteristic of the BEKK parameterisation is its parsimonious specification and the presence of paired transposed matrices guarantees the symmetry and the non-negativeness of \( H_t \), with no additional a-priori restrictions. The BEKK framework also reduces the number of parameters to be estimated without imposing strong constraints on the shape of the interaction between markets.
A fifth advantage of this specification is the easy introduction of a leverage term. In terms of the number of parameters to be estimated for a BEKK model that accounts for leverage effects, the variance system requires 11 parameters to be estimated for two equations, 24 parameters for three equations and 42 parameters for four equations. The parameters cannot be interpreted on an individual basis. Instead, the functions of the parameters that form the intercept terms and the coefficients of the lagged variance, covariance and residual terms are of interest.

3.4.3 Estimation Procedure

Since the GARCH model is no longer of the usual linear form, ordinary least squares cannot be used for parameter estimation. There are a variety of reasons for this - the simplest and most fundamental is the ordinary least squares procedure minimises the residual sum of squares. In turn, the residual sum of squares depends only on the parameters in the conditional mean equation and not on the conditional variance. Therefore, residual sum of squares minimisation is no longer an appropriate objective. This section discusses a maximum-likelihood technique that may be employed in estimating nonlinear models such as GARCH.

The likelihood function is a multiplicative function of the observed data, that consequently is difficult to maximise with respect to the parameters. Therefore, its logarithm is taken in order to transform the likelihood function into an additive function of the sample data - the log-likelihood function. The log-likelihood function is maximised in the analysis using the RATS package.\(^{38}\)

\(^{38}\)The conditional mean equation is equal to \(\Delta y_t = \mu + e_t\), where \(y_t\) is the natural logarithm of
In maximising the log-likelihood function, RATS employs an iterative procedure. Given a set of initial estimates for the conditional variance parameters, the estimates are updated at each iteration until the program determines that an optimum has been reached. Normally distributed errors are assumed in the estimation process, implying the following log-likelihood function may be expressed as

\[ L(\theta) = -\frac{N}{2}ln(2\pi) - \frac{1}{2}\sum_{t=1}^{N}(ln|H_t| + e_t'H_t^{-1}e_t), \]  

(3.9)

where \( \theta \) represents the vector of parameters to be estimated, \( N \) is the number of observations and \( H_t \) is the \( 2 \times 2 \) time-varying conditional variance-covariance matrix.

If the log-likelihood function has only one maximum with respect to the parameter values, any optimisation method should be able to find this maximum. As is often the case with nonlinear models such as GARCH, the log-likelihood function may have many local maxima, so that different algorithms may find different local maxima of the log-likelihood function. The optimisation method employed by RATS is based on the determination of the first and (estimated) second derivatives of the likelihood function with respect to the parameter values at each iteration, known as the gradient and Hessian respectively.\(^{39}\)

The algorithm implemented for the nonlinear optimisation of the log-likelihood function is the BFGS algorithm described by Broyden, Fletcher, Goldfarb and Shanno (Broyden, 1970; Fletcher, 1970; Goldfarb, 1970; Shanno, 1970), implemented via the RATS package. The BFGS algorithm attempts to find a local min-

---

\(^{39}\)The Hessian is the matrix of second derivatives of the log-likelihood function with respect to the parameters.
imum of the objective function by taking steps in the error space whose directions are determined based on the partial derivatives of the given objective function.

Since the assumption of conditional normality cannot be maintained, robust estimates of the covariance matrices of the parameter estimates are calculated using the BFGS algorithm. Under fairly weak conditions, the resulting estimates are also consistent when the conditional distribution of the residuals is non-normal (Bollerslev and Wooldridge, 1992).

Before the BFGS algorithm is implemented, a technique known as the simplex algorithm may be employed. The simplex algorithm is a derivative-free technique since it does not make use of the score vector in determining the direction of the following iteration. The method in maximising the log-likelihood function in the thesis involves initially using the simplex method (for 30 iterations), followed by the BFGS algorithm.

An issue in estimating GARCH models is the choice of initial values for the parameters. The standard method for achieving these starting values is to estimate a linear regression model for the mean equation and to set the estimates of the conditional mean parameters from the linear regression as initial values for the nonlinear optimisation. The residual vector is usually initiated using the residuals from the linear regression estimation, and each element of the conditional variance $H_t$ is initialised by setting it equal to the unconditional variance of the ordinary

---

40Since it does not use derivative methods, the simplex method is usually slower to reach an optimum, and cannot produce standard error estimates. The simplex method is more robust to local curvature than the BFGS algorithm.
least squares residuals. The other parameters in the conditional variance equation must be set somewhat arbitrarily. By default, any parameters that do not have a starting value are assigned by RATS as having an initial value of zero. A zero value is often a local optimum for the log-likelihood function, making non-zero starting values plausible.

This section reviewed a maximum-likelihood procedure that is implemented in the estimation of nonlinear (GARCH) models in this thesis. The log-likelihood function is maximised on the assumption of normally distributed errors via an iterative procedure using the RATS package. The next step is to extend the foundation (GARCH) model to include information provided by cointegration-based information. The GARCH-X specification is now briefly discussed.

3.5 The GARCH-X Model

Due to the potential for spurious regression, it is not valid to perform ordinary least squares regression on non-stationary variables. If, for example, the variables are all trending upwards then any correlation between them may simply reflect this fact rather than a genuine correlation between the variables. Usually researchers would like to test if a relationship exists between variables that are non-stationary. This is achieved by testing whether or not the variables are cointegrated, that is, if the variables have a common trend (Granger, 1981; Engle and Granger, 1987; Johansen, 1988; Johansen and Juselius, 1990; Layton and Tan, 1992).

The notion of cointegration - via the error correction term - may be linked to
the GARCH model to provide a specification that accounts for short-run deviations from the long-run cointegration relationship. The squared spread between spot and futures rates may have potential predictive power in modelling volatility of asset returns, volatility that is not captured effectively by the GARCH (1,1) model. Hedge ratios and hedging performance may be less than optimal when the phenomena of cointegration is mistakenly omitted from the hedging model. This section extends the GARCH model to incorporate cointegration information.

Lee (1994) modelled the conditional heteroscedasticity of the prediction error of foreign exchange rates via the implementation of an extension of the GARCH model (to include the error correction term of the cointegrated series). The model is known as the GARCH-X model. Lee (1994) found that since the conditional variance of the prediction error is positively related to the squared spread, unmodelled conditional heteroscedasticity in the GARCH(1,1) model may be explained by a function of the spread - the squared value of the error correction term lagged once.

The idea behind the novel approach of Lee (1994) revolves around the error correction model being considered as a prediction equation of the cointegrated series (being a conditional expectation of the first difference of the series given an information set) and is logical when one expects increased volatility due to shocks to the system that propagate on both the first and second moments (Lee, 1994). The error correction term imposes a long-run relationship between spot and futures prices. In the presence of a deviation, either the futures price is too high (low) and/or the spot price is too low (high). To correct the imbalance, the spot price increases (decreases) and/or the futures price decreases (increases).
The BEKK formulation of the GARCH-X model is given by

\[ H_t = C'C + A'e_{t-1}e'_{t-1}A + B'H_{t-1}B + D'Dz^2_{t-1}, \]  

(3.10)

where D is a 1 × 2 matrix of coefficients and z_t is the error correction term equalling \( S_t - \beta F_t \) (where \( S_t \) and \( F_t \) - the natural logarithms of the spot and futures prices respectively - are cointegrated with cointegration parameter \( \beta \)). When long-run equilibrium exists, the GARCH-X model abstracts information from the error correction term \( (S_t - \beta F_t) \), so the size of the deviations (or the variance of the deviations) from the long-run equilibrium level may provide added information about the relationship between spot and futures rates. A direct implication is the further the deviations from long-run equilibrium, the more volatile the asset prices and the more difficult they are to forecast. Such added information may be exploited to obtain more precise time-varying confidence intervals for point forecasts of asset returns.

In the GARCH-X model, the squared error is determined by the information set available at time t, denoted by \( F_{t-1} \). Therefore, the conditional variance \( H_t \) may be written as a measurable function of the variables in \( F_{t-1} \). The estimates of the coefficients of the error correction term may quantify the extent that the model explains the relationship between disequilibrium and conditional volatility not explained by the GARCH(1,1) specification. Significance of the coefficients of the error correction term implies that last period’s equilibrium error has significant impact on the adjustment process of the subsequent returns. Therefore, deviations in the short-run in one period are not adjusted in the subsequent period. In regards to hedging, this implies that short-run deviations have substantial impact on the
optimal hedge ratio.

The more influential the cointegration relationship - where spot and futures
prices are more responsive to the deviation from the equilibrium level - the more
costly it may be for a hedger to omit the cointegration variable in the statistical
modelling of spot and futures prices. The GARCH-X model may possibly be superior
to the GARCH model due to the ability to capture and adequately model such
disequilibriums, thus providing a better model of the conditional heteroscedasticity
of asset prices.

The GARCH and GARCH-X methods may not be of benefit in efficient markets
since there should not be substantial room for improvement in these markets, as
basis risk is small due to the volatility usually being known and not clustered. Even
when the volatility is clustered, there is usually a pattern to it that is known by
participants in the market, who account for the clustering when finding a “fair value”
of the asset or commodity.

The GARCH-X model may be applied to determine whether any additional
information, above and beyond that obtained via a GARCH model alone, may be
extracted about:

1. the prediction of future short-run covariances.

2. the dynamics of the short-run correction back to the average expected covari-
   ance.

There would appear to be at least three possible ways in which the GARCH-X
model may perhaps contribute to the modelling of hedging (in particular to cross-
hedging). The model may provide:

1. information pertaining to the calculation of the number of derivative (futures, forwards, options etc.) contracts and their respective duration.

2. insight into the manner in which the hedges may be dynamically altered as the time series paths of asset prices are traced out.

3. information about the choice of cross-hedging asset or commodity.

As has been mentioned, there appears to be certain factors that mitigate against the efficiency (or effectiveness) of hedging. These conditions have to hold in order to utilise the information from a cointegration perspective. Potentially, cointegration does provide added information. In cross-hedging situations, cointegration may prove to be beneficial since the volatilities are unknown, not random and clustered. If spot and futures prices are not cointegrated, the hedged portfolio suffers from the risk of potentially large changes in its value.

### 3.6 Conclusion

This chapter outlined various dynamic specifications for the conditional second moments. The distributional characteristics of asset returns were highlighted along with various tests for determining the extent of any possible departures from normality. The ARCH/GARCH family of statistical models was discussed since the examples in the analysis involve the application of GARCH and GARCH-X models. After introducing the ARCH model, the GARCH specification was outlined both in a uni-
variate and multivariate framework. Such models are well-designed to investigate the persistence of volatility in the trend and the transitory components. In particular, two specifications of the MGARCH model - the VEC and BEKK formulations - are outlined prior to application of the BEKK formulation of the GARCH model in later examples. The chapter also outlined the method in which the parameters in a GARCH model are estimated in the analysis.

This chapter discussed an extension of the GARCH formulation - known as the GARCH-X specification (Lee, 1994) - in modelling the distributional properties of basis spreads. The analysis focussed on determining whether the improved modelling of the basis creates more effective hedging protection relative to models that are less effective in capturing these distributional properties. The GARCH-X model incorporates cointegration information into the GARCH process in dealing with time-dependent basis risk variance. The GARCH-X specification is potentially theoretically superior to other existent formulations since it retains information on long-run behaviour of spot and futures prices as well as providing information pertaining to the short-term fluctuations from the long-term equilibrium level.
Chapter 4

Determination of Dynamic Hedging Procedures

4.1 Introduction

The previous chapter outlined several alternative formulations for the conditional second moments of asset returns. The thesis concentrates on the BEKK specification of the bivariate GARCH and GARCH-X models. In this chapter, the BEKK GARCH formulation is applied to model the variance-covariance matrix of spot and futures returns. Dynamic hedge ratios are then constructed and compared to their static alternatives - in terms of their comparative ability to minimise the risk of the hedged portfolio.

The purpose of this chapter is to:

1. develop criteria to determine under which conditions different hedging strategies are more effective from the perspective of a bona-fide hedger. Optimality
in this instance is defined solely in terms of risk-reduction. The ability of
dynamic approaches to minimise risk is compared to both the conventional
static procedures of naive and minimum-variance hedging. If the hedge ratios
are unstable, allowance for such stochastic movements is shown to substan-
tially increase hedging effectiveness by reducing the volatility of the hedged
portfolio.

2. propose an alternative constant hedge to the minimum-variance and naive
techniques. The innovative hedge is termed the “forecasted hedge” and is
based upon the forecasting curves of the conditional covariance of spot and
futures returns and the conditional variance of the futures returns. The “fore-
casted hedge” ratio provides a constant alternative to the minimum-variance
hedge. Hedging criteria is determined that enables a comparison of two con-
stant hedge ratios, bypassing the need for transaction cost considerations since
portfolio re-balancing costs are now redundant. The minimum-variance hedge
is found to not necessarily provide the most effective constant hedge. An indi-
cation of the preferred constant hedge is provided by examining the conditional
variances of the constructed hedged portfolios for the out-of-sample period.

The motivation behind the research is to determine the conditions under which a
given hedge ratio is more effective - in terms of greater risk-reduction - than another
hedge ratio. A hedge ratio based on the limit of the forecasting curve of future
hedge ratios may provide a more effective alternative static hedge to the minimum-
variance procedure, depending upon appropriately modelling the within-sample data
This chapter proceeds as follows: Section 4.2 highlights the limitations of constant hedges such as the minimum-variance hedge. Section 4.3 compares various hedging procedures via the formulation of conditional variance comparisons. Section 4.4 introduces the “forecasted hedge” - a constant hedge calculation based on the conditional covariance and conditional variance forecasts. Section 4.5 compares the naive, minimum-variance and “forecasted hedge” procedures. Section 4.6 summarises the issues raised in the chapter and the conclusions reached, and identifies the analytical steps that follow.

4.2 Shortcomings of Conventional Hedging Models

This section documents the limitations of conventional hedging procedures. The shortcomings of traditional models motivate consideration of dynamic techniques in constructing, and maintaining, effective hedge ratios. The major limitation of conventional (static) hedging models is discussed, providing a link to the following section that compares variance measures between different hedging techniques.

Since the research work of Mandelbrot (1963) and Fama (1965), researchers have known that volatility in the covariance between spot and futures returns, and the volatility in both variances, is clustered. The conditional distribution of such returns may be a heavy-tailed non-Gaussian distribution generating outliers. Although the returns themselves are uncorrelated, their squared values are highly correlated in the
sense that the squared return in a prior period and the subsequent period squared return are likely to be very similar in value - behaviour that is indicative of nonlinear data generating processes (Hsieh, 1991, 1995).

The minimum-variance hedging rule is ineffective where the basis is not constant since the slope parameter from the minimum-variance approach is merely a ratio of the unconditional covariance between spot and futures returns to the unconditional variance of the futures returns. However, where the basis is time varying, the covariance and variance in the optimal hedging rule are clearly conditional moments that depend upon information available at the time the hedging decision is made. In these situations, a preferred alternative to static hedging is a generalised approach that takes proper account of relevant conditioning information. Improvements to hedging performance may be made through the implementation of strategies that involve more than simply buying and holding a fixed futures position over the entire cash-holding period.

When constructing hedges, consideration should be made for possible time-variation in the distribution of spot and futures returns and the hedge ratio should be continuously re-balanced at appropriate time intervals. However, a dynamic procedure may not be beneficial since the associated costs of implementing a dynamic technique are even higher due to the constant updating of the hedge ratio. Drawbacks of dynamic hedging approaches such as cost and complexity must therefore be weighed up against the improvement such techniques provide. While dynamic models are likely to possess superior explanatory power (for the movement of returns) with respect to the static model, such increased power does not automatically trans-
late into dynamic techniques being cost/benefit superior to static procedures, except in periods of increased volatility (Watt, 1997). In such situations, static approaches fail to provide adequate hedging performance by failing to take into account the dynamic interaction between spot and futures returns. However, Watt (1997) concluded that a static hedging strategy may indeed be the optimal cost/benefit hedging procedure even though dynamic models may have superior explanatory power.

The impact of hedging on basis risk in underdeveloped markets is a crucial issue in emerging economies. The intrinsic structural factors in immature (or non-existent) futures markets may make effective hedging difficult due to thin trading and associated increases in premiums and price spreads. Lack of market accessibility may further exacerbate the problem of adequate spot/risk coverage. The basis possesses a time-dimension because the correlation between spot and futures returns may not be as strong as when direct hedging is invoked - in the latter situation the correlation between spot and futures returns may be expected to be stable because of arbitrage-free pricing relationships. However, assuming constant correlation may not be appropriate in cross-hedging because spot and futures returns might not change on a one-to-one basis since these returns represent different, albeit related, assets. Therefore, the resultant hedging strategy should allow correlations and variances to change in light of new information.

41Direct hedging is riskless when one hedges a security with a futures contract in the same security, the hedge is lifted at contract expiration and interest rates are deterministic. No such assurance exists for a cross-hedge. If lifted at any time, both hedges are subject to “basis risk”.
4.3 Comparison of Hedging Methods

This section develops criteria to determine the situations under which each of the naive, minimum-variance and dynamic hedging methods are the most effective techniques. Each pair of hedging procedures (naive, minimum-variance, forecasted hedge and time-varying) are compared. Although the practical usefulness of a hedging procedure can best be assessed where effective hedge ratios determined in a prior period can be implemented in a subsequent period, the criteria developed here allows for a hedge to be updated if the current hedge ratios are found to lose their effectiveness over time. Discrepancies between ex-post and ex-ante hedging effectiveness raises concerns over the practicality and reliability of the hedge for risk-reduction. Optimality for the bona-fide hedger aims to minimise the variance of the returns ex-ante.

Traditional comparison of hedging methods is based on the calculation of the portfolio return at each out-of-sample period $t$, where the daily portfolio return is defined by Equation (2.5). The (unconditional) variance of these constructed returns is calculated over the out-of-sample (forecasting) period and the hedging method that leads to the smallest variance of portfolio returns is regarded as the most effective hedging procedure.\textsuperscript{42} Within this time-frame it may be possible to ascertain certain periods where various hedging alternatives are likely to provide more effective hedges. Conditional variance measures are formed and analysed, the logic behind this is that a certain hedging method may be most effective during

\textsuperscript{42}The criterion may only be applied when the quantity $\Delta S_{t-1} - h_t \Delta F_{t-1}$ is stable. An estimate of the variance of portfolio returns may be obtained from the sample variance of the returns.
certain periods but not during others.

The criteria whereby the naive and minimum-variance hedge ratios are equally effective as the time-varying hedge ratio are derived and, more importantly, from this criteria conditions may be determined under which this dynamic hedge is superior to static hedges. The first task involves the calculation of the conditional variance of each of the naive, minimum-variance and time-varying hedging techniques.

4.3.1 Conditional Variance of Different Hedging Strategies

In the case of naive hedging, \( h_t = 1 \) for all \( t \). Therefore, the conditional variance of such a hedged portfolio is given by

\[
\text{Var}_{t-1}(\Delta S_t - \Delta F_t) = H_{11,t} - 2H_{12,t} + H_{22,t},
\]

where \( H_{11,t} \) denotes the conditional variance of spot returns at time \( t \), \( H_{12,t} \) denotes the conditional covariance between spot and futures returns at time \( t \), and \( H_{22,t} \) denotes the conditional variance of futures returns at time \( t \).

The minimum-variance hedge ratio, denoted by \( h_{MV} \), is given by

\[
h_{MV} = (X'X)^{-1}(X'Y),
\]

where

\[
Y = \begin{pmatrix}
\Delta S_1 \\
\vdots \\
\Delta S_{t-1}
\end{pmatrix}, \quad X = \begin{pmatrix}
\Delta F_1 \\
\vdots \\
\Delta F_{t-1}
\end{pmatrix}
\]

Therefore,

\[
h_{MV} = \frac{\sum_{i=1}^{t-1} \Delta S_i \Delta F_i}{\sum_{i=1}^{t-1} \Delta F_i^2},
\]
and the conditional variance of a portfolio hedged by the minimum-variance procedure is given by

\[ Var_{t-1}(\Delta S_t - h_{MV} \Delta F_t) = H_{11,t} - 2h_{MV} H_{12,t} + h_{MV}^2 H_{22,t} \]

Where the hedge is constructed via time-varying techniques, the hedge ratio is defined by \( h_t = \frac{H_{12,t}}{H_{22,t}} \). The conditional variance of the resultant hedged portfolio is

\[ Var_{t-1}(\Delta S_t - h_t \Delta F_t) = H_{11,t} - 2h_t H_{12,t} + h_t^2 H_{22,t} \]

\[ = H_{11,t} - 2 \frac{H_{12,t}}{H_{22,t}} + \frac{H_{12,t}^2}{H_{22,t}} \]

\[ = H_{11,t} - \frac{H_{12,t}^2}{H_{22,t}} \quad (4.1) \]

The quantity on the right-hand side of Equation (4.1) is smaller as \( H_{12,t} \) approaches \( H_{11,t} H_{22,t} \). Therefore, the portfolio variance decreases as the conditional (dynamic) correlation between spot and futures returns increases. The correlation between spot and futures returns, according to Hegde (1982), is expected to rise during periods of increased volatility.

### 4.3.2 Comparisons Between Different Hedging Strategies

First, the minimum-variance procedure is compared to the naive technique. The minimum-variance procedure provides greater risk-reduction than the naive technique when

\[ Var_{t-1}(\Delta S_t - h_{MV} \Delta F_t) - Var_{t-1}(\Delta S_t - \Delta F_t) < 0, \]

that is, when

\[ H_{11,t} - 2h_{MV} H_{12,t} + h_{MV}^2 H_{22,t} - H_{11,t} + 2H_{12,t} - H_{22,t} < 0, \]

122
which may be simplified to

\[(h_{MV} - 1)((h_{MV} + 1)H_{22,t} - 2H_{12,t}) < 0\]

Since the hedge ratio is usually assumed to lie between 0 and 1, the minimum-variance hedge may be accepted as being more effective than the naive hedge when

\[(h_{MV} + 1)H_{22,t} - 2H_{12,t} > 0\]

Second, the time-varying procedure is compared to the naive technique. The dynamic hedge is assumed to produce a more effective hedge - in terms of risk-reduction - than the naive technique when

\[Var_{t-1}(\Delta S_t - h_t \Delta F_t) - Var_{t-1}(\Delta S_t - \Delta F_t) < 0,\]

that is, when

\[H_{11,t} - \frac{H_{12,t}^2}{H_{22,t}} - H_{11,t} + 2H_{12,t} - H_{22,t} < 0,\]

which may be simplified to

\[-(H_{12,t} - H_{22,t})^2 < 0,\]  \hspace{1cm} (4.2)

as the quantity \(H_{22,t}\) is strictly positive for all \(t\). The left-hand side of Equation (4.2) is always strictly non-positive, therefore the dynamic hedge is never less effective than the naive hedge. This is based on the assumption that \(H_{11,t}, H_{12,t}\) and \(H_{22,t}\) are effectively modelled through time. In the case where \(H_{12,t} \approx H_{22,t}\) for all \(t\) - when the time-varying hedge ratio is approximately equal to 1 for all \(t\), implying perfect correlation - both the dynamic and naive techniques are equally effective, otherwise the time-varying hedge is more effective than the naive hedge.
Lastly, the time-varying procedure and the minimum-variance technique are compared. The dynamic technique provides better risk-reduction than the minimum-variance hedge when

\[ \text{Var}_{t-1}(\Delta S_t - h_t \Delta F_t) - \text{Var}_{t-1}(\Delta S_t - h \Delta F_t) < 0, \]

that is, when

\[ H_{11,t} - \frac{H_{12,t}^2}{H_{22,t}} - H_{11,t} + 2h_{MV}H_{12,t} - h_{MV}^2H_{22,t} < 0, \]

which may be simplified to

\[ -\left( \frac{H_{12,t}}{H_{22,t}} - h_{MV} \right)^2 < 0, \quad (4.3) \]

since the quantity \( H_{22,t} \) is strictly positive for all \( t \). Since the left-hand side of Equation (4.3) is always non-positive, the time-varying hedge is never less effective than the minimum-variance hedge. The minimum-variance hedge is only equally effective as the dynamic hedge when the quantity \( \frac{H_{12,t}}{H_{22,t}} \) is approximately constant for all \( t \).

The above comparisons suggest that naive hedges are adequate when the conditional ratio \( \frac{H_{12,t}}{H_{22,t}} \) approximates 1 over the interval \( t_0 \) to \( t_1 \). Alternatively, if the ratio \( \frac{H_{12,t}}{H_{22,t}} \) is approximately equal to a constant (not necessarily 1), a minimum-variance hedge is effective over the period from \( t_0 \) to \( t_1 \). In the third case of unstable (time-varying) hedge ratios, dynamic techniques should be applied in hedging a desired spot position. In this case, a plot of the quantity \( \frac{H_{12,t}}{H_{22,t}} \) is likely to reveal non-stationarity and thus the most effective hedge does not remain constant during the life of the hedge. The comparable effectiveness of the dynamic hedge with respect
to each constant hedging rule depends upon the degree of instability in the most effective hedge ratio $\frac{H_{12,t}}{H_{22,t}}$ over time.

### 4.3.3 Comparison Between Constant Hedge Ratios

This section will compare the effectiveness of two constant hedge ratios and the mechanics behind such a comparison. Assume there is the choice of two available hedge ratios, $h_1$ and $h_2$, both of which are constant. Furthermore, assume that $h_1 > h_2$. Under such conditions, $h_1$ is determined to be superior to $h_2$ when

$$\text{Var}_{t-1}(\Delta S_t - h_1 \Delta F_t) - \text{Var}_{t-1}(\Delta S_t - h_2 \Delta F_t) < 0,$$

that is,

$$H_{11,t} - 2h_1H_{12,t} + h_1^2H_{22,t} - H_{11,t} + 2h_2H_{12,t} - h_2^2H_{22,t} < 0,$$

which is equivalent to

$$-2h_1H_{12,t} + h_1^2H_{22,t} + 2h_2H_{12,t} - h_2^2H_{22,t} < 0,$$

and simplifies to

$$(h_1 - h_2)((h_1 + h_2)H_{22,t} - 2H_{12,t}) < 0$$

Therefore, the acceptance of $h_1$ as a superior hedge ratio to $h_2$ depends upon whether the quantity $(h_1 + h_2)H_{22,t} - 2H_{12,t}$ is less than zero. Since $h_t = \frac{H_{12,t}}{H_{22,t}}$, the rule may be simplified to determining whether $\frac{h_1 + h_2}{2} < h_t$. If so, $h_1$ is accepted to be more effective than $h_2$. Intuitively this appears logical since the most effective hedge ratio at time $t$ is $h_t$. 125
In empirical situations, the hedge ratio cannot be altered frequently due to transaction costs. Therefore, between two constant hedge ratios, \( h_1 \) and \( h_2 \), one should choose the hedge ratio that is closer to \( h_t \) for the majority of time periods. Alternatively, when there appears to be certain periods of time in which one hedge ratio is more effective than the other, this hedging rule may be invoked and may lead to an application of different hedge ratios over different periods of time.

In the next section the forecasted hedge is proposed, calculated as the limit of the forecasting curve of the dynamic hedge ratio (available at present time \( t \)). The forecasted hedge is then compared to both the constant hedge ratio from the minimum-variance technique and the naive hedge ratio.

4.4 Forecasted Hedge Ratios

This section provides a generalisation of the \( f \)-step ahead forecast of the conditional variance \( H_{t+f} \) - available at time \( t \). The generation provides an alternative constant hedge ratio to the naive and minimum-variance techniques and is labelled the forecasted hedge. The advantage of such a hedge ratio is that dynamic information may be extracted from \( H_{t+f} \) to provide a hedge ratio that is static. In turn, the hedging strategy may involve a switching of hedge ratios (between the minimum-variance and the forecasted hedges, and vice-versa).

Although the hedge ratios themselves cannot be forecast explicitly, the conditional variance matrix may be forecast \( f \)-steps ahead and the resultant hedge ratio may be obtained in the conventional manner. The minimum-variance approach
is one such static procedure that determines the hedge ratio from past data and employs this ratio to out-of-sample observations.

The next task is to illustrate how the forecasted hedge ratio can be derived. The forecasted hedge is defined as the limit as \( f \) approaches infinity (\( \infty \)) of the forecast of the conditional covariance of spot and futures returns \( (H_{12,t+f}) \) divided by the forecast of the conditional variance of the futures returns \( (H_{22,t+f}) \). In practice, the ratio \( \frac{H_{12,t+f}}{H_{22,t+f}} \) may stabilise over a relatively short period of time.

Assume the bivariate GARCH specification is represented by Equation (3.8). At present time \( t \), the resulting \( f \)-step ahead forecast is given by

\[
\hat{H}_{t+f} = E_t(e_{t+f}e'_{t+f})
\]

Defining \( A \), \( B \) and \( C \) as matrices in the GARCH specification, the one-step ahead forecast is given by

\[
\hat{H}_{t+1} = E_t(e_{t+1}e'_{t+1}) = C'C + A'e_tA + B'H_tB
\]

The task in this section is to derive a general expression for \( \hat{H}_{t+f} \). These \( f \)-step ahead forecasts have been used to obtain the forecasting curve of the covariance matrix that in turn is utilised to obtain the forecasted hedge ratio. Since the covariance matrix has a limiting value for each component, the forecasted hedge ratio is defined as the limit as \( f \) approaches \( \infty \) of the forecasted covariance of spot and futures returns divided by the forecasted variance of the futures returns. The general form of the \( f \)-step ahead forecast is now provided.
Definition 1 Given matrices \( A, B \) and \( D \), and assuming \( A'DA + B'DB \) is well-defined, the operator \( \otimes \) is defined as
\[
(A' + B') \otimes D \otimes (A + B) = A'DA + B'DB
\]

Under Definition 1, given matrices \( A, B \) and \( C \),
\[
(A' + B') \otimes C'C \otimes (A + B) = A'C'CA + B'C'CB,
\]
providing the quantity \( A'C'CA + B'C'CB \) is well-defined, and
\[
(A' + B') \otimes (A' + B') \otimes C'C \otimes (A + B) \otimes (A + B)
\]
\[
= (A' + B') \otimes D \otimes (A + B)
\]
\[
= A'DA + B'DB,
\]
where
\[
D = (A' + B') \otimes C'C \otimes (A + B)
\]
\[
= A'C'CA + B'C'CB,
\]
if both \( A'DA + B'DB \) and \( A'C'CA + B'C'CB \) are well-defined. Define
\[
\bigotimes_{i=1}^{l} (A' + B') = (A' + B') \otimes \cdots \otimes (A' + B'),
\]
The next task is to show the general form of the \( f \)-step ahead forecast of the conditional variance is given by
\[
\hat{H}_{t+f} = E_t(e_{t+f}e_{t+f}')
\]
\[
= C'C + \sum_{i=1}^{f-1} \left( \bigotimes_{i=1}^{l} (A' + B') \right) \otimes C'C \otimes \left( \bigotimes_{i=1}^{l} (A + B) \right)
\]
\[
+ \left( \bigotimes_{i=1}^{f-1} (A' + B') \right) \otimes A'e_{t+i}A \otimes \left( \bigotimes_{i=1}^{f-1} (A + B) \right)
\]
\[
+ \left( \bigotimes_{i=1}^{f-1} (A' + B') \right) \otimes (B'H_iB) \otimes \left( \bigotimes_{i=1}^{f-1} (A + B) \right)
\]
For $f = 1$,

$$
\hat{H}_{t+1} = E_t(e_{t+1}e'_{t+1})
$$

$$
= C'C + A'e_t e'_t A + B'H_t B
$$

Therefore, the one-step ahead forecast of the hedge ratio is given by

$$
\hat{H}_{t+1} = C'C + A'e_t e'_t A + B'H_t B.
$$

If $f = 2$,

$$
\hat{H}_{t+2} = E_t(e_{t+2}e'_{t+2})
$$

$$
= C'C + (A' + B') \bigotimes C'C \bigotimes (A + B)
$$

$$
+ (A' + B') \bigotimes (A'e_t e'_t A) \bigotimes (A + B)
$$

$$
+ (A' + B') \bigotimes (B'H_t B) \bigotimes (A + B)
$$

$$
= C'C + A'C'CA + B'C'CB + A'^2e_t e'_t A^2
$$

$$
+ B'A'e_t e'_t AB + A'B'H_t BA + B'^2H_t B^2
$$

Therefore, the two-step ahead forecast of the hedge ratio is given by

$$
\hat{H}_{t+2} = C'C + A'C'CA + B'C'CB + A'^2e_t e'_t A^2 + B'A'e_t e'_t AB
$$

$$
+ A'B'H_t BA + B'^2H_t B^2.
$$

For $f = 3$,

$$
\hat{H}_{t+3} = E_t(e_{t+3}e'_{t+3})
$$

$$
= C'C + \sum_{i=1}^{2} \left( \bigotimes_{i=1}^{l} (A' + B') \bigotimes C'C \bigotimes \left( \bigotimes_{i=1}^{l} (A + B) \right) \right)
$$

$$
+ \left( \bigotimes_{i=1}^{2} (A' + B') \bigotimes (A'e_t e'_t A) \bigotimes \left( \bigotimes_{i=1}^{2} (A + B) \right) \right)
$$

129
\[ + \left( \bigotimes_{i=1}^{2} (A' + B') \right) \bigotimes (B'H_t B) \bigotimes \left( \bigotimes_{i=1}^{2} (A + B) \right) \]
\[ = C'C + (A' + B') \bigotimes C'C \bigotimes (A + B) \]
\[ + (A' + B') \bigotimes (A' + B') \bigotimes C'C \bigotimes (A + B) \bigotimes (A + B) \]
\[ + (A' + B') \bigotimes (A' + B') \bigotimes (A'e_t e'_t A) \bigotimes (A + B) \bigotimes (A + B) \]
\[ + (A' + B') \bigotimes (A' + B') \bigotimes (B'H_t B) \bigotimes (A + B) \bigotimes (A + B) \]
\[ = C'C + A'C'CA + B'C'CB \]
\[ + (A' + B') \bigotimes (A'C'CA + B'C'CB) \bigotimes (A + B) \]
\[ + (A' + B') \bigotimes (A'^2 e_t e'_t A^2 + B'A'e_t e'_t AB) \bigotimes (A + B) \]
\[ + (A' + B') \bigotimes (A'B'H_t BA + B'^2 H_t B^2) \bigotimes (A + B) \]
\[ = C'C + A'C'CA + B'C'CB + A'^2 C'CA^2 + A'B'C'CBA \]
\[ + B'A'C'cab + B'^2 C'CB^2 + A'^6 e_t e'_t A^3 + A'B'A'e_t e'_t ABA \]
\[ + B'A'^2 e_t e'_t A^2 B + B'^2 A'e_t e'_t AB^2 + A'^2 B'H_t BA^2 + A'B'^2 H_t B^2 A \]
\[ + B'A'B'H_t BAB + B'^3 H_t B^3 \]

Therefore, the three-step ahead forecast of the hedge ratio is given by
\[
\hat{H}_{t+3} = C'C + A'C'CA + B'C'CB + A'^2 C'CA^2 + A'B'C'CBA
+ B'A'C'cab + B'^2 C'CB^2 + A'^6 e_t e'_t A^3 + A'B'A'e_t e'_t ABA + B'A'^2 e_t e'_t A^2 B
+ B'^2 A'e_t e'_t AB^2 + A'^2 B'H_t BA^2 + A'B'^2 H_t B^2 A + B'A'B'H_t BAB + B'^3 H_t B^3.
\]

The forecasts generated by this notation agree with the original notation, namely
\[
\hat{H}_{t+1} = E_t(e_{t+1} e'_{t+1})
= C'C + A'e_t e'_t A + B'H_t B
\]
\[
\hat{H}_{t+2} = E_t(e_{t+2} e'_{t+2})
\]
\[
\hat{H}_{t+3} = E_t(e_{t+3}e'_t) \\
= E_t(E_{t+1}(e_{t+2}e'_{t+2})) \\
= E_t(C'C + A'e_{t+1}e'_{t+1}A + B'H'_{t+1}B) \\
= C'C + A'(C'C + A'e_te_tA + B'H_B)A \\
+ B(C'C + A'e_te_tA + B'H_B)B \\
= C'C + A'C'CA + A'^2e_te_tA^2 + A'B'H_BA \\
+ B'C'CB + B'A'e_te_tA + B'^2H_B^2, \\
\]

\[
\hat{H}_{t+3} = E_t(e_{t+3}e'_t) \\
= E_t(E_{t+1}(E_{t+2}(e_{t+3}e'_t))) \\
= E_t(E_{t+1}(C'C + A'e_{t+2}e'_{t+2}A + B'H_{t+2}B)) \\
= E_t(C'C + A'(C'C + A'e_{t+1}e'_{t+1}H_{t+1}A + B'H_{t+1}B)A \\
+ B'(C'C + A'e_{t+1}e'_{t+1}A + B'H_{t+1}B)B \\
= C'C + A'C'CA + A'^2(C'C + A'e_te_tA + B'H_B)A^2 \\
+ A'B'(C'C + A'e_te_tA + B'H_B)BA + B'C'CB \\
+ B'A'(C'C + A'e_te_tA + B'H_B)AB \\
+ B'^2(C'C + A'e_te_tA + B'H_B)B^2 \\
= C'C + A'C'CA + A'^2C'CA^2 + A'^3e_te_tA^3 + A'^2B'H_BA^2 \\
+ A'B'C'BA + A'B'A'e_te_tABA + A'B^2H_B^2A + B'C'CB \\
+ B'A'C'AB + B'A'^2e_te_tA^2B + B'A'B'H_BAB \\
+ B'^2C'CB^2 + B'^2A'e_te_tA^2B + B'^3H_B^3. 
\]
By induction, it can be shown the $f$-step ahead forecast of the conditional variance, $\hat{H}_{t+f}$, can be expressed as

$$\hat{H}_{t+f} = C'C + \sum_{i=1}^{f-1} \left\{ \bigotimes_{l=1}^{l} (A' + B') \right\} \bigotimes C'C \bigotimes \left\{ \bigotimes_{i=1}^{f-1} (A + B) \right\}$$

$$+ \left[ \bigotimes_{i=1}^{f-1} (A' + B') \right] \bigotimes A'e_iA \bigotimes \left[ \bigotimes_{i=1}^{f-1} (A + B) \right]$$

$$+ \left[ \bigotimes_{i=1}^{f-1} (A' + B') \right] \bigotimes (B' \hat{H}B) \bigotimes \left[ \bigotimes_{i=1}^{f-1} (A + B) \right]$$

This section derived the $f$-step ahead forecast of the conditional variance $H_{t+f}$, available at time $t$. The forecasted hedge provides an alternative constant hedge ratio that can act as a surrogate for time-varying conditional variance models. The forecasted hedge extracts dynamic information from $H_{t+f}$ to provide a hedge ratio that is static. The forecasted hedge is applied in the following section.

4.5 Applications of Dynamic Hedging Procedures

To gain an understanding of the potential strength of the forecasted hedge ratio, two examples are shown. In these examples the hedging effectiveness is compared between four types of hedges:

1. the naive hedge.
2. the minimum-variance hedge.
3. the forecasted hedge.
4. the time-varying (GARCH) hedge (where this dynamic hedge is re-balanced daily).
By re-balancing daily the dynamic hedge provides the most effective hedge ratio, $h_t$, based upon information available at time $t$. The following analysis involves three tasks:

1. testing for non-normality such as skewness, leptokurtosis and volatility clustering in spot and futures returns, determining the significance of any deviation from normality.

2. choosing the specification of the conditional mean and conditional second moment equations.

3. constructing the GARCH hedge ratios and subsequent application of these hedge ratios to spot and futures returns, obtaining values for the effectiveness of various hedging techniques.

4.5.1 Application to Simulated Data

Example 1

In this example, two time series of returns ($\Delta S_t$ and $\Delta F_t$) are generated from the GARCH model via implementation of the S-Plus statistical package. The forecasted hedge ratio is determined via the method outlined in the previous section using information available at time $t$. Daily returns are constructed as implied by the computed hedge ratios and the variance of the returns of the constructed portfolios are calculated over the entire sample period to determine whether the application of the GARCH model results in increased hedging effectiveness when compared
to static hedging procedures. The forecasted hedge ratio is compared to the two alternative static techniques.

Data is simulated from a bivariate GARCH model. The sample size is chosen to be 1000, the last 100 observations of which are with-held to comprise the out-of-sample (forecasting) period. The parameters in the GARCH specification are

\[
A = \begin{pmatrix} 0.4 & 0 \\ 0 & 0.4 \end{pmatrix} \\
B = \begin{pmatrix} 0.9 & 0 \\ 0 & 0.9 \end{pmatrix} \\
C = \begin{pmatrix} 0.02 & 0 \\ 0.02 & 0.025 \end{pmatrix}
\]

Since \(A_{11}^2 + B_{11}^2 = 0.97 < 1\) and \(A_{22}^2 + B_{22}^2 = 0.97 < 1\) in the above GARCH model, current shocks to either the spot or futures return series do not persist indefinitely in conditioning the future variance of the returns (Kavussanos and Nomikos, 2000a, 2000b). Both sets of returns are therefore covariance stationary. Since both \(A_{11}^2 + B_{11}^2\) and \(A_{22}^2 + B_{22}^2\) are close to unity, there is significant (though not infinite) persistence.

In the GARCH model the half-life for the spot returns is equal to the half-life for the futures returns, which in turn is equal to 23.76 days.\(^{43}\)

\(^{43}\)The half-life may be defined as the length of time (in days) taken for the initial impact on the market of a shock to reduce to half its original size (Kavussanos and Nomikos, 2000a). The formula for calculating the half-life is \(1 - \frac{\ln(2)}{\ln(A_{ii}^2 + B_{ii}^2)}\), for \(i = 1, 2\), where \(\ln\) denotes the natural logarithm. The closer to unity the value of \(A_{ii}^2 + B_{ii}^2\), the slower the decay rate and the longer the half-life.
The time series plots of the returns, $\Delta S_t$ and $\Delta F_t$, appear in Figures 4.1 and 4.2 respectively. The measures of skewness of $\Delta S_t$ and $\Delta F_t$ are both significant (at the 1% level of significance). Similarly, the measure of kurtosis of $\Delta S_t$ is significant at the 1% level of significance. However, the kurtosis present in $\Delta F_t$ is not significant. The Jarque-Bera test reveals that both $\Delta S_t$ and $\Delta F_t$ are not normally distributed (at the 1% level of significance).

The Ljung-Box test reveals that both return series are autocorrelated (at the 1% level of significance for the spot returns and at the 5% level of significance for the futures returns). The absolute values of $\Delta S_t$ and $\Delta F_t$ and $(\Delta S_t)^2$ and $(\Delta F_t)^2$ exhibit similar autocorrelation (all at the 1% level of significance). The signs for the skewness of spot and futures returns are negative and positive respectively, indicating the distributions of each series are skewed left and right respectively. The Lagrange-Multiplier test (at the 1% level of significance) confirms the autocorrelation in both return series (ARCH effects). Table 4.1 reports the summary statistics.

The true parameters are adopted in this example to eliminate any potential errors that may result from the estimation of the unknown parameters of the GARCH model. The time series plot of the GARCH hedge ratios for the out-of-sample period is shown in Figure 4.3. The GARCH hedge ratios for the out-of-sample (forecasting) period vary from 0.3992 to 0.8369 with a mean value of 0.6248.

The minimum-variance hedge ratio equals 0.3513 within-sample, this hedge ratio underestimates the correlation between spot and futures returns in all out-of-sample periods. Similarly, the forecasted hedge yields a hedge ratio of 0.7889, a higher ratio than that of the minimum-variance technique, though lower than the naive
Figure 4.1: Time series plot of the spot returns for Example 1.

Figure 4.2: Time series plot of the futures returns for Example 1.
Table 4.1: Summary statistics for the spot and futures returns for Example 1.

<table>
<thead>
<tr>
<th></th>
<th>$\Delta S_t$</th>
<th>$\Delta F_t$</th>
<th>5% Critical Values</th>
<th>1% Critical Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>900</td>
<td>900</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.784$^a$</td>
<td>0.404$^a$</td>
<td>(-0.160,0.160)</td>
<td>(-0.210,0.210)</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>4.693$^a$</td>
<td>2.860</td>
<td>(2.680,3.320)</td>
<td>(2.580,3.420)</td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>917.895$^a$</td>
<td>331.230$^a$</td>
<td>5.991</td>
<td>9.210</td>
</tr>
<tr>
<td>Ljung-Box $Q(24)$</td>
<td>44.256$^a$</td>
<td>40.756$^b$</td>
<td>36.415</td>
<td>42.980</td>
</tr>
<tr>
<td>Ljung-Box $</td>
<td>Q(24)</td>
<td>$</td>
<td>1797.516$^a$</td>
<td>804.936$^a$</td>
</tr>
<tr>
<td>Ljung-Box $Q^2(24)$</td>
<td>1531.815$^a$</td>
<td>1013.230$^a$</td>
<td>36.415</td>
<td>42.980</td>
</tr>
<tr>
<td>Lagrange-Multiplier</td>
<td>314.830$^a$</td>
<td>256.767$^a$</td>
<td>36.415</td>
<td>42.980</td>
</tr>
</tbody>
</table>

Notes:
1. $a$ denotes significance at the 1% level of significance.
2. $b$ denotes significance at the 5% level of significance.

Figure 4.3: Time series plot of the hedge ratios obtained via implementation of the bi-variate GARCH model for Example 1.
hedge ratio. The GARCH hedge ratios in Figure 4.3 reveal that the forecasted hedge (0.7889) is closer to the GARCH hedge ratios than the minimum-variance hedge ratio (0.3513) for the majority of time periods. The plots of the differences between the conditional variances for all combinations of hedging techniques appear in Figures 4.4-4.9.

The hedging effectiveness value of the GARCH model is equal to 0.9007. The hedging effectiveness of the naive, minimum-variance and forecasted hedges are 0.6619, 0.7013 and 0.8646 respectively. The most effective constant hedge is the forecasted hedge, the hedging effectiveness of which is very close to that of the GARCH hedging technique and nearly 25% more effective than the minimum-variance hedge. The major benefit of the forecasted hedge is the ratio remains constant throughout the out-of-sample (forecasting) period, thus avoiding the re-balancing costs that result from the frequent re-balancing of the hedge portfolio.

4.5.2 Application to Financial Data

Example 2

This example focusses on empirical data, obtained from the Turtle – Trader website (www.turtletrader.com).\(^{44}\) The data analysed involve the closing prices for heating oil futures and crude oil futures contracts, traded on the New York Mercantile Exchange (NYMEX).\(^{45}\) The data commence on January 4, 1993 and

\(^{44}\)Turtle – Trader offers free downloads of historical futures data for numerous currencies and commodities.

\(^{45}\)The NYMEX website (www.nymex.com) contains contract specifications for various types of
Figure 4.4: Conditional variance of the time-varying (GARCH) hedge minus the conditional variance of the minimum-variance hedge for Example 1.

Figure 4.5: Conditional variance of the time-varying (GARCH) hedge minus the conditional variance of the naive hedge for Example 1.
Figure 4.6: Conditional variance of the time-varying (GARCH) hedge minus the conditional variance of the forecasted hedge for Example 1.

Figure 4.7: Conditional variance of the minimum-variance hedge minus the conditional variance of the naive hedge for Example 1.
Figure 4.8: Conditional variance of the minimum-variance hedge minus the conditional variance of the forecasted hedge for Example 1.

Figure 4.9: Conditional variance of the naive hedge minus the conditional variance of the forecasted hedge for Example 1.
terminate on March 31, 1998. The modelling period finishes on December 31, 1997 and the period from January 2, 1998 to March 31, 1998 is assumed to be the out-of-sample (forecasting) period. The out-of-sample period consists of 61 observations while the sampling period consists of 1252 observations.

The hedger is assumed to have a long position in the heating oil futures contract and wishes to hedge the exposure to this contract by shorting the crude oil futures contract. In the analysis, the natural logarithms of the heating oil futures prices are denoted by $Heat$ while the natural logarithms of the crude oil futures prices are denoted by $Crude$. The hedger would like to cover the exposure to $Heat$ using $Crude$. The time series plots of the prices (natural logarithms) of the heating oil and crude oil futures contracts, as well as the time series plots of the returns for both the heating oil and crude oil futures contracts, are shown in Figures 4.10-4.13.

From Figures 4.12 and 4.13 it may be seen that both return series are time-varying. Furthermore, $\Delta Heat$ and $\Delta Crude$ are found to be non-normal. Skewness and kurtosis for both returns are significant at the 1% level. The returns are skewed left and are heavy-tailed. The Jarque-Bera test reveals significant non-normality in both returns series. The Ljung-Box and Lagrange-Multiplier tests show the returns are autocorrelated. The tests are all significant at the 1% level of significance, with the exception of the Ljung-Box test for $\Delta H_t$, which is significant at the 5% level. Both returns series therefore are characterised by ARCH effects. The summary

---

46The analysis is equally applicable if the hedger has a short position in the crude oil futures contract and wishes to hedge the exposure to this contract by shorting the heating oil futures contract.
Figure 4.10: *Time series plot of the heating oil futures prices for Example 2.*

Figure 4.11: *Time series plot of the crude oil futures prices for Example 2.*
Figure 4.12: Time series plot of the heating oil futures returns for Example 2.

Figure 4.13: Time series plot of the crude oil futures returns for Example 2.
statistics are included in Table 4.2.

Table 4.2: Summary statistics for the heating oil futures returns and crude oil futures returns in Example 2.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>( \Delta H_t )</th>
<th>( \Delta C_t )</th>
<th>5% Critical Values</th>
<th>1% Critical Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>1252</td>
<td>1252</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Skewness</td>
<td>-1.044(^a)</td>
<td>-0.378(^a)</td>
<td>(-0.152,0.152)</td>
<td>(-0.200,0.200)</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>6.449(^a)</td>
<td>5.200(^a)</td>
<td>(2.696,3.304)</td>
<td>(2.601,3.399)</td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>2396.841(^a)</td>
<td>1440.258(^a)</td>
<td>5.991</td>
<td>9.210</td>
</tr>
<tr>
<td>Ljung-Box ( Q(24) )</td>
<td>41.962(^b)</td>
<td>46.483(^a)</td>
<td>36.415</td>
<td>42.980</td>
</tr>
<tr>
<td>Ljung-Box (</td>
<td>Q(24)</td>
<td>)</td>
<td>342.664(^a)</td>
<td>188.460(^a)</td>
</tr>
<tr>
<td>Ljung-Box ( Q^2(24) )</td>
<td>170.831(^a)</td>
<td>158.520(^a)</td>
<td>36.415</td>
<td>42.980</td>
</tr>
<tr>
<td>Lagrange-Multiplier</td>
<td>112.791(^a)</td>
<td>105.909(^a)</td>
<td>36.415</td>
<td>42.980</td>
</tr>
</tbody>
</table>

Notes:
1. \(^a\) denotes significance at the 1% level of significance.
2. \(^b\) denotes significance at the 5% level of significance.

The bivariate specification applied to model the returns of *Heat* and *Crude* is given by

\[
\Delta y_t = \mu + e_t, \quad (4.4)
\]

where

\[
y_t = \begin{pmatrix} \text{Heat}_t \\ \text{Crude}_t \end{pmatrix}, \quad e_t = \begin{pmatrix} e_{\text{Heat},t} \\ e_{\text{Crude},t} \end{pmatrix}
\]

and \( e_t | \mathcal{F}_{t-1} \sim N(0, H_t) \). The conditional structure of the covariance matrix \( H_t \) is based on the BEKK specification outlined in Equation (3.8).

As the model is nonlinear, numerical methods are required to estimate the parameters. Under the assumption the random errors are normally distributed, the
BFGS optimisation algorithm outlined in Chapter 3 is implemented to obtain estimates (and the corresponding standard errors) of the parameters in the bivariate GARCH(1,1) model. The estimates are obtained by maximising the nonlinear log-likelihood function in Equation (3.9), where $H_t$ is the $2 \times 2$ time-varying conditional variance-covariance matrix. The estimates of the GARCH coefficients within-sample (that is, from January 4, 1993 to December 31, 1997) are calculated via the econometric package RATS and are included in Table 4.3.

Table 4.3: Parameter estimates obtained via implementation of the bivariate GARCH model for Example 2.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_1$</td>
<td>-0.0003(^a)</td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>-0.0001(^a)</td>
</tr>
<tr>
<td>$C_{11}$</td>
<td>0.0025(^a)</td>
</tr>
<tr>
<td>$C_{12}$</td>
<td>0.0023(^a)</td>
</tr>
<tr>
<td>$C_{22}$</td>
<td>0.0017(^a)</td>
</tr>
<tr>
<td>$A_{11}$</td>
<td>0.2764(^a)</td>
</tr>
<tr>
<td>$A_{22}$</td>
<td>0.2306(^a)</td>
</tr>
<tr>
<td>$B_{11}$</td>
<td>0.9543(^a)</td>
</tr>
<tr>
<td>$B_{22}$</td>
<td>0.9629(^a)</td>
</tr>
<tr>
<td>Log-likelihood Function Value</td>
<td>9380.77</td>
</tr>
</tbody>
</table>

Notes:
1. \(^a\) denotes significance at the 1% level of significance.
2. \(^b\) denotes significance at the 5% level of significance.

Current shocks to either the spot or futures returns series do not infinitely persist in conditioning the future variance of the returns (since $A_{11}^2 + B_{11}^2 = 0.9870 < 1$ and $A_{22}^2 + B_{22}^2 = 0.9804 < 1$) - both sets of returns are covariance stationary. However,
as $A_{11}^2 + B_{11}^2$ and $A_{22}^2 + B_{22}^2$ are both close to unity, there is significant persistence.

In this example, the half-life for the spot returns is equal to 54 days, while the half-life for the futures returns is equal to 36 days. The hedge ratios obtained through application of the GARCH model to the out-of-sample (forecasting) period vary from 0.5880 to 0.9503, with a mean value of 0.8133. The time series plot of the GARCH hedge ratios for the out-of-sample period is shown in Figure 4.14.

Figure 4.14: *Time series plot of the hedge ratios obtained via implementation of the bivariate GARCH model for Example 2.*

The within-sample minimum-variance hedge ratio equals 0.6297. The minimum-variance hedge ratio underestimates the correlation between spot and futures returns in nearly all out-of-sample periods. Similarly, the forecasted hedge ratio equals 0.7563, a higher ratio than the minimum-variance technique, though lower than the naive hedge ratio. The forecasted hedge leads to increased hedging effectiveness.
over the minimum-variance technique. The GARCH hedge ratios in Figure 4.14 are, in the main, closer to the forecasted hedge (0.7563) than the minimum-variance hedge ratio (0.6297) for the majority of time periods. The differences between the conditional variances for all combinations of hedging techniques appear in Figures 4.15-4.20.

The hedging effectiveness value when the GARCH model is applied is equal to 0.9385. The hedging effectiveness measures of the naive, minimum-variance and forecasted hedges are 0.8816, 0.9054 and 0.9434 respectively. The forecasted hedge provides the most effective hedge. The hedger may also obtain a hedging effectiveness value of 0.9299 by alternating between the naive and minimum-variance hedges according to the criteria (that is, conditional variance comparisons) outlined in this chapter. This hedging effectiveness measure may be obtained by using the minimum-variance hedge for the first 16 time periods, followed by the naive hedge from period 17 to 40 inclusively and then reverting back to the minimum-variance hedge from period 41 onwards. Important information is provided by the GARCH model, allowing the hedger to alternate between naive and minimum-variance hedges.

4.6 Conclusion

Traditional hedging techniques make an over-simplified assumption of a constant variance-covariance matrix of spot and futures returns, leading to less effective hedging decisions than conditional hedges. In order to achieve more effective hedges, allowance should be made for the possibility that the variances of spot and futures
Figure 4.15: Conditional variance of the time-varying (GARCH) hedge minus the conditional variance of the minimum-variance hedge for Example 2.

Figure 4.16: Conditional variance of the time-varying (GARCH) hedge minus the conditional variance of the naive hedge for Example 2.
Figure 4.17: Conditional variance of the time-varying (GARCH) hedge minus the conditional variance of the forecasted hedge for Example 2.

Figure 4.18: Conditional variance of the minimum-variance hedge minus the conditional variance of the naive hedge for Example 2.
Figure 4.19: Conditional variance of the minimum-variance hedge minus the conditional variance of the forecasted hedge for Example 2.

Figure 4.20: Conditional variance of the naive hedge minus the conditional variance of the forecasted hedge for Example 2.
returns and/or the covariances between these returns exhibit dynamic characteristics. If added information is available about the relationship between spot and futures returns, the basic hedge may then be re-adjusted by buying or selling more contracts with variations in basis risk. Dynamic adjustments should preserve hedging effectiveness.

This chapter developed statistical criteria based on conditional variance measures to permit comparison of the effectiveness of various hedge ratios. The analysis is applicable to either a bona-fide or a relatively risk-averse hedger, or where the futures prices are martingales. The ability of dynamic approaches to minimise risk is compared to the ability of various conventional static hedge ratios. Where the hedge ratios are unstable, allowance for such stochastic movements substantially increases hedging effectiveness by reducing the volatility of the hedged portfolio.

The minimum-variance hedge does not necessarily provide the most effective constant hedge from the perspective of a bona-fide hedger interested only in minimising risk. An alternative constant hedge, known as the “forecasted hedge”, was proposed. The “forecasted hedge” is based upon the forecasting curves of the conditional covariance of spot and futures returns and the conditional variance of the futures returns. The “forecasted hedge” may provide greater risk-reduction than any other static hedge and may be a viable hedging alternative if the investor does not wish to adopt a time-varying hedge. The significance of the “forecasted hedge” is that it extracts information from the parameters estimated from the GARCH model and applies this information to obtain forecasting curves for both the variances of spot and futures returns and the covariances between these returns. The “fore-
casted hedge” may be compared to any other constant hedge through the adoption of hedging rules. Further variance reduction may be obtained by switching (constant hedge) regimes from the minimum-variance (forecasted) procedure to the forecasted (minimum-variance) technique.
Chapter 5

The Concept of Cointegration

5.1 Introduction

The previous chapter outlined the applicability of dynamic hedging techniques to both simulated and empirical data. The analysis is limited solely to the BEKK specification of the bivariate GARCH model. The GARCH-X model, as outlined in Chapter 3, is applied to both empirical and simulated data in Chapter 6. The GARCH-X model is an extension of the GARCH model that allows the incorporation of cointegration information into the modelling of the conditional second moments. This chapter explains the concept of cointegration along with its various definitions.

The analysis in Chapter 4 showed that GARCH models are the most promising category of time series models for effectively incorporating characteristics of financial asset time series that lead to non-normality in the residual processes - such as leptokurtosis and heteroscedasticity. GARCH models are therefore used in the current investigation dealing with conditional volatility between two or more finan-
cial time series. As mentioned, GARCH models have limitations. Specifically, they omit information on long-run behaviour and provide little information about the short-run shock adjustment process.

The incorporation of cointegration analysis directly into the GARCH model appears to be a promising method of overcoming these limitations. The process of incorporation commences in this chapter by explaining those fundamentals of cointegration relevant to this task. The traditional Johansen cointegration technique is not entirely appropriate for this task because the Johansen procedure assumes that a finite order VAR model with white noise is an adequate proxy to a finite moving-average process. Therefore, the Johansen method may be inaccurate in the presence of moving-average errors within a time series.

The purpose of this chapter is to:


2. explain the theory and procedure of an alternative technique developed in Lin and McCrae (1999, 2001a) - the residual-based cointegration (RBC) procedure - that overcomes the limitations of traditional cointegration techniques.

3. illustrate the application of the RBC procedure to simulated and empirical data.

The chapter is organised as follows: Section 5.2 introduces the concept of cointegration and the benefits it provides in modelling long-run equilibrium. Section 5.3
outlines the traditional definitions of cointegration, section 5.4 introduces an alternative approach to the Johansen procedure. The fundamental link between cointegration vectors and residual processes obtained from the fitting of univariate ARIMA models is explained, motivating the RBC technique. The section also describes the tasks in the practical application of the RBC approach. Section 5.5 analyses examples of autoregressive and moving-average series to show the applicability of the RBC procedure. Section 5.6 investigates the accuracy of the cointegration vector estimate for simulated series drawn from a model example in Section 5.5. Section 5.7 contains empirical applications of the RBC approach to foreign exchange data. Section 5.8 summarises the potential contribution of the RBC approach in financial time series analysis and suggests possible avenues for future research.

5.2 A Measure of Long-Run Equilibrium

Prior to the formal development of cointegration techniques, the usual practice in univariate time series model fitting procedures was to difference the data to achieve stationarity. However, differencing implies the loss of long-run information. A GARCH model that includes cointegration enables the simultaneous investigation into both the long-run relationship and short-run dynamic adjustments to such a long-run relationship. This section introduces the concept of cointegration and its benefits as a measure of long-run equilibrium.

The long-run equilibrating potential of financial time series such as foreign exchange rates, stock prices and interest rates, and the relationship of this potential
to market efficiency in both spot and forward markets is probably the most intensively researched topic in cointegration analysis (Copeland, 1991; Crowder, 1996). Cointegration is said to occur when attractor forces within a vector of financial time series keep the series in close proximity or “long-run equilibrium” so that a linear combination of component series forms a stationary series in itself (Granger, 1981; Engle and Granger, 1987; Johansen, 1988; Johansen and Juselius, 1990; Layton and Tan, 1992). Should the variables drift away from equilibrium for a certain period of time, equilibrating economic forces will restore equilibrium.

The concept of cointegration, introduced by Granger (1981) and further developed by Engle and Granger (1987), defined long-run equilibrium between elements of a time series vector as the existence of a linear combination of these vector elements that is stationary. Granger (1986) argued the presence of cointegration between two speculative markets for two different assets implies predictability and hence, market inefficiency. Hakkio and Rush (1989) concluded that market efficiency implies that even when spot and forward rates are non-stationary, they will be cointegrated.

Cointegration enables the incorporation of non-stationarity, long-term relations and short-run properties into the one modelling process (Engle and Yoo, 1987), thereby overcoming several limitations of classical univariate inference analysis. The notion of cointegration gives a statistically precise definition to the concept of a long-run equilibrium process. The importance of such a precise definition is two-fold:

1. Cointegration provides a framework for empirical testing of “equilibrium” theory in finance and economics.
2. Cointegration allows the modelling of long-run behaviour of financial asset pricing time series, especially in the area of derivative markets.

Cointegration enables the determination of whether or not asset prices are characterised by long-run equilibrating factors. An absence of these factors implies that asset price time series are completely random (non-deterministic) in the long-run as well as in the short-run. Finance theory would say that in the short-run the price is non-deterministic because the series are supposedly random. The Efficient Market Hypothesis states that at any point, asset prices efficiently incorporate all available information about present and future expectations. Consequently the direction and magnitude of future asset price movements is indeterminate. Economic theory says that while this may be true in the short-term, over longer periods of time one would expect movements towards a long-run equilibrium. Information about the long-run behaviour of cointegrated time series allows precise testing of the presence (or otherwise) of long-run equilibrating factors between series in a cointegrated vector.

A problem in testing market efficiency is that financial price series are generally non-stationary. In such cases, conventional statistical procedures are no longer appropriate because they tend to bias toward incorrectly rejecting efficiency. Unlike these procedures, cointegration does not require stationarity in eligible time series and so avoids this bias as well as preserving long-run behaviour. If two price series are cointegrated, equilibrating factors create a long-run stationary process. If the price series are not cointegrated, they will tend to deviate apart without bound, which is contrary to the Efficient Market Hypothesis. However, when the spot prices in two different markets are cointegrated, one of the prices must help in forecasting
the other, providing an arbitrage opportunity.

In terms of trying to predict the underlying future asset price, cointegration information may not be beneficial, since the asset price is random or may be slightly deterministic. However, due to the arbitrage-free relationship, the basis may help determine the price movements and should therefore be considered in modelling the long-run equilibrium variance. At time $t$, one may proceed forward to time $t + k$ and conclude there are a number of possibilities where basis risk might be at that point in time (since spot and futures prices are known to be cointegrated and the covariance is time-dependent). There are three such possibilities: the conditional variance is either on the long-run equilibrium variance, is below the long-run equilibrium variance or is above it.

Cointegration not only provides the answer as to which of the above three possible scenarios occur but, where the spread between series is mean-reverting, may provide information in regards to the expected path and speed at which the conditional variance reverts to the long-run equilibrium variance. Therefore, while in terms of period forward forecasting cointegration is likely to be of little use, it may be beneficial as a forecasting tool because it provides information about the expected behaviour of the present variance in terms of the cointegration variance. Forecast enhancement from exploiting cointegration comes from using information in the current deviations from the cointegration relationships. Therefore, knowing whether and by how much the cointegration relations are violated today is valuable in assessing where the variables are tomorrow.
5.3 Traditional Procedures of Cointegration

This section defines the concepts of “cointegration”, “order of integration” and “bi-variate” and “multivariate” cointegration. These definitions will then be used to describe the traditional definitions of cointegration. The definitions are readily available from Harris (1995).

**Definition 2** A time series $X_t$, that has a stationary ARMA representation after differencing $d$ times but with the term $(1-L)^{d-1}X_t$ being non-stationary, is integrated of order $d$ and is denoted by $X_t \sim I(d)$, where $L$ is the back-shift (lag) operator.

**Definition 3** Let $X_t = (X_{1,t}, \cdots, X_{p,t})'$ be a time series vector, $t = 1, 2, \ldots, n$. If each component of $X_t$ is $I(1)$ and there exists a vector $\xi$ such that $\xi'X_t \sim I(0)$, $X_{1,t}, \cdots, X_{p,t}$ are said to be cointegrated and $\xi$ is called a cointegration vector for the system $X_t$.

An $I(0)$ series may be defined as a stationary, trend-free series. An $I(1)$ series is such that its first difference is $I(0)$. These two types of series have quite different appearances and properties. In particular, under reasonable assumptions, the variance of an $I(0)$ series is bounded whereas the unconditional variance of an $I(1)$ series increases without bound as $t$ increases.

5.3.1 The Engle-Granger Approach to Cointegration

Engle and Granger (1987) introduced the technique of cointegration. Generally stated, two variables are said to be cointegrated and share a common stochastic trend
if they move together for a long period of time. More formally, two variables that are $I(1)$ are cointegrated if there exists a stationary linear combination between them.

To test for cointegration in the bivariate case, one has to only find statistical evidence the residuals of a linear combination of the two variables are stationary. If two variables are cointegrated, causality from $X$ and $Y$ can be tested by either examining the coefficient of the one-period lagged error term of the cointegration equation of the two variables or by examining the significance of the lagged differences of the variable $X$ (Engle and Granger, 1987).

An important factor in formulating time series models is the stability of the component series. For instance, Phillips (1987) finds that spurious regressions often occur in a set of time series variables that are $I(1)$ but not cointegrated. To avoid this problem, any model expressed in variable levels must be restricted to cointegrated variables only. Where the variables are not cointegrated, any model that involves non-cointegrated variables should be stated in their first differences. Furthermore, for any cointegrated system there may be more than one cointegration vector. Stock and Watson (1988) show that if $n$ variables share $n - r$ ($r < n$) common non-stationary trends, the variables contain $r$ cointegration vectors.

The first test for cointegration has been ascribed to Engle and Granger (1987). Their cointegration analysis tests for a long-term linear relationship between economic variables. The first step in the procedure is to determine the order of integration of each series by applying unit root tests. The tests presume the order of differencing is an integer. The second task involves the differencing of series until they are all of the same order - conventionally $I(1)$. The third step is to test whether
the error term in the cointegration regression is $I(0)$. An $I(0)$ error term implies the error term exhibits \textit{mean – reverting} behaviour and there exists a long-run equilibrium relationship between the series in the cointegration regression. Alternatively, an $I(1)$ error term implies non-stationarity and an absence of any long-run equilibrium relationship between the series in the cointegration regression.

### 5.3.2 The Johansen Approach to Cointegration

The Engle and Granger (1987) technique only allows identification of a single cointegration vector within a system. Johansen (1988) derived a procedure that overcomes this limitation by being able to identify multiple linearly independent cointegration vectors within a system. In the Johansen method, the evolution of eligible time series is assumed to be well-specified by a finite order VAR model with white noise. The Johansen method will now be outlined mathematically to simplify the understanding of the fundamentals behind the procedure.

Letting $X_t$ be a $(p \times 1)$ vector of $I(1)$ variables - variables that become stationary after taking first differences -

\begin{equation}
X_t = \Pi_1 X_{t-1} + \Pi_2 X_{t-2} + ... + \Pi_k X_{t-k} + e_t, \ t = 1, 2, ..., T, \tag{5.1}
\end{equation}

where each of the $\Pi_k$ is a $(p \times p)$ matrix of parameters and $e_t \sim IN(0, \Sigma)$.

Equation (5.1) may be transformed into

\[ \Delta X_t = \Gamma_1 \Delta X_{t-1} + ... + \Gamma_{k-1} \Delta X_{t-k+1} + \Pi X_{t-k} + e_t, \]

where $\Gamma_i = -(1 - \Pi_1 - ... - \Pi_i), \ i = 1, 2, ..., k - 1$ and $\Pi = -(1 - \Pi_1 - ... - \Pi_k)$. 

162
The matrix $\Pi$ contains information about the long-run properties of the model. If $\Pi$ has rank 0, the system is not cointegrated (all the variables in $X_t$ are integrated of order one or higher). If $\Pi$ has rank $p$ (full rank), the variables in $X_t$ are stationary. If $\Pi$ has rank $r$ (where $0 < r < p$), $\Pi$ may be decomposed into two distinct ($p \times r$) matrices $\alpha$ and $\beta$ such that $\Pi = \alpha \beta'$. In this case, there are $r$ independent cointegration vectors, given by $\beta$. The parameter $\alpha$ represents the speed of adjustment to disequilibrium and $\beta$ is a matrix of long-run coefficients. If there are $r \leq (n - 1)$ independent cointegration vectors in $\beta$, this implies the last $(p - r)$ columns of $\alpha$ are zero. Therefore, the determination of how many $r \leq (p - 1)$ cointegration vectors exist in $\beta$ is equivalent to testing which columns of $\alpha$ are zero.

The Johansen method overcame several disadvantages of the Engle-Granger procedure associated with multiple cointegration vectors, limiting distributions and test robustness. The ability of Johansen’s maximum likelihood technique to identify all potential multiple cointegration vectors within a given vector of eligible time series made this method much more attractive. The Johansen method is more powerful than the Engle-Granger method as the former provides test statistics for the number of cointegration vectors that have an exact limiting distribution, which is a function of only one parameter (Hall, 1989). The maximum likelihood estimates of the unconstrained cointegration vectors are not dependent on any arbitrary normalisation. Johansen (1988) demonstrates that likelihood ratio tests have asymptotic distributions that are a function only of the difference between the number of variables and the number of cointegration vectors. Johansen’s method is also more robust and less susceptible to small sample bias than the Engle-Granger approach.
Unlike the Engle-Granger approach, the Johansen method provides a unified approach to estimation and testing of cointegration relations in the framework of vector autoregression (VAR) error correction models. There are three further advantages of the Johansen procedure:

1. the Johansen method estimates the cointegration vector within the context of a complete error correction model. The use of the complete model provides empirical estimates of each of the cointegration vectors.

2. the Johansen procedure provides estimates of the response dynamics between variables in terms of the time taken for the system to reach a long-run equilibrium (speed of adjustment) following shocks to the system.

3. the Johansen method tests for restrictions on individual elements of the cointegration vector(s) and those restrictions imposed on the parameters in the model.

A main limitation to the Johansen approach is its inappropriateness in the presence of moving-average errors within a time series. The Johansen technique assumes the underlying data generating process is well-fitted by a finite order VAR model with white noise. The presence of a moving-average error violates this assumption. Therefore, Johansen’s technique presumes that empirically a finite order VAR with white noise proves an adequate proxy to a finite moving-average process. The assumption may represent a severe practical limitation, especially with “chaotic” financial time series. Forcing the data into a mis-specified model to satisfy these
conditions reduces accuracy in estimating cointegration vector coefficients in comparison with fitting a more appropriate model.

5.4 An Alternative Approach to Cointegration

The limitation of the Johansen procedure in the presence of moving-average errors is overcome in an approach developed by Lin and McCrae (1999, 2001a). This alternative procedure for determining cointegration vectors, when the components of the system of interest satisfy certain conditions, allows the underlying time series to be fitted by ARIMA models other than finite order VAR’s. Using this innovative procedure, the significance of the cointegration behaviours may be identified and estimated easily. This section outlines the RBC approach to cointegration and the tasks involved in implementing such a procedure.

Although the idea is similar to that of Engle and Yoo (1987) and Bierens (1997), the difference between the procedure of Lin and McCrae (1999) and that of both Engle and Yoo (1987) and Bierens (1997) is that the latter approaches focus their attention on standard VAR models where the covariance matrix has full rank. The novelty of the RBC approach involves an analysis of the relationship between multivariate cointegration and univariate ARIMA modelling where each individual time series is modelled independently. In such univariate modelling, investigation only requires analysis of the covariances between individual series. Furthermore, the covariance matrix of the residual processes given by the ARIMA models need not be of full rank.
The RBC approach utilises the theoretical relationship that exists between the Engle-Granger and Johansen cointegration procedures and univariate ARIMA model fitting techniques for individual time series. The approach provides a method for examining:

1. whether a relationship between cointegration and univariate ARIMA modelling exists.

2. whether the cointegration vector(s) for the system may be determined by univariate ARIMA model fitting procedures, or more specifically, the covariance matrix of the residual processes.

Under certain weak conditions, Lin and McCrae (1999, 2001a) show that the number of cointegration vectors may be determined via the rank of the covariance matrix of the residual processes. This is done as follows: given a system \( \mathbf{X}_t = (X_{1,t}, \ldots, X_{p,t})' \), assume that all of the elements of \( \mathbf{X}_t \) are \( I(1) \). If \( \mathbf{X}_t \) can be accepted as a cointegrated system, the following procedure may be applied to determine all linearly independent cointegration vectors. The procedure consists of five tasks:

1. Fit \( \mathbf{X}_t \) by an appropriate ARIMA model, say

\[
\Phi(B)(1 - B)\mathbf{X}_t = \mu + \Theta(B)e_t, \quad t = 1, 2, \ldots, n
\]

where \( e_t = (e_{1,t}, \ldots, e_{p,t})' \) are white noise, \( \mu = (\mu_1, \ldots, \mu_p)' \) and

\[
\Phi(B) = \begin{pmatrix}
\Phi_1(B) & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
0 & 0 & \ddots & \ddots \\
0 & 0 & \cdots & \Phi_p(B)
\end{pmatrix}, \quad \Theta(B) = \begin{pmatrix}
\Theta_1(B) & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
0 & 0 & \ddots & \ddots \\
0 & 0 & \cdots & \Theta_p(B)
\end{pmatrix},
\]
where all the $\Phi_i(B)$ and $\Theta_i(B)$ are finite order polynomial functions of $B$ with roots outside the unit circle.

2. If the residual vectors $\mathbf{e}_t = (e_{1,t}, \ldots, e_{p,t})'$ and residual cross-products $e_{i,t}e_{j,t}$, $i < j \leq p$, $t = 1, 2, \ldots, n$ are both stationary (or have ergodic properties) the residual vectors may be analysed to obtain the sample covariance matrix for $\mathbf{e}_t$. The sample covariance matrix is denoted by $\hat{\Sigma}_n$, and is implemented to estimate $\mathbf{e}_t$.

3. Determine the eigenvalues of $\hat{\Sigma}_n$, say $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{p-r} > \lambda_{p-r+1} \geq \ldots \geq \lambda_p$, and corresponding eigenvectors. The eigenvectors form a matrix denoted by $\mathbf{A}$. The matrix $\mathbf{A}$ may be re-written as $\mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2)$, where $\mathbf{A}_2$ is formed by those eigenvectors corresponding to the smaller eigenvalues $\lambda_{p-r+1} \geq \lambda_{p-r+2} \geq \ldots \geq \lambda_p$.

4. Let $v^{(1)}_t = (v_{1,t}, \ldots, v_{p-r,t})'$ and $v^{(2)}_t = (v_{p-r+1,t}, \ldots, v_{p,t})'$ satisfy the following equation

$$
\Phi(B)(1 - B)\mathbf{X}_t = \mu + \Theta(B)\mathbf{e}_t, \quad t = 1, 2, \ldots, n,
$$

where

$$
\mathbf{e}_t = (\mathbf{A}_1 \mathbf{A}_2) \begin{pmatrix} v^{(1)}_t \\ v^{(2)}_t \end{pmatrix}
$$

Now $\mathbf{X}_t$ may be expressed as

$$
(1 - B)\mathbf{X}_t = \bar{\mu} + C(B)v^{(1)}_t + C_1(B)v^{(2)}_t,
$$
with

\[ C(B) = \Phi(B)^{-1} \Theta(B) A_1 \]

\[ C_1(B) = \Phi(B)^{-1} \Theta(B) A_2, \quad \text{and} \]

\[ \tilde{\mu} = \Phi(B)^{-1} \mu \]

Therefore,

\[ X_t = \tilde{\mu} t + X_0 + (1 - B) \frac{C(B) - C(1)}{1 - B} \sum_{i=1}^{t} v_i^{(1)} \]
\[ + C(1) \sum_{i=1}^{t} v_i^{(1)} + C_1(B) \sum_{i=1}^{t} v_i^{(2)}, \]

where \( C(1) = C(B) \) with \( B = 1 \). Upon solving for \( \xi' C(1) = 0 \), \( \xi \) is obtained such that

\[ \xi' (X_t - \tilde{\mu}) = \xi' X_0 - \xi' W_0 + \xi' W_t + \xi' C_1(B) \sum_{i=1}^{t} v_i^{(2)}, \]

where

\[ W_t = \frac{C(B) - C(1)}{1 - B} v_t^{(1)}, \]

and the term

\[ \xi' X_0 - \xi' W_0 + \xi' W_t \]

is \( I(0) \) (Lin and McCrae, 1999).

Therefore, when the impact of the non-stationary component - \( \xi' C_1(B) \sum_{i=1}^{t} v_i^{(1)} \) - is not significant, \( \xi \) can be accepted as a cointegration vector for \( X_t - \tilde{\mu} t \). An important issue is how to determine whether the non-stationary component is in fact significant. The size of the impact of this non-stationary component
may be measured by the ratio of variances given by the non-stationary and stationary components in the manner of Lin and McCrae (2001b).

5. Verify the stationarity of the linear combination of the cointegration vector $\xi'(X_t - \tilde{\mu}t)$ by application of the Augmented Dickey-Fuller (Dickey and Fuller, 1981) (ADF) test and by graphical procedures.

5.5 Application of the RBC Procedure to Simulated Data

This section outlines how the method of RBC may be applied in practice via the consideration of simulated data. In theory this alternative approach is theoretically valid but a sub-issue is whether this translates into practical significance. There are two stages in the application of the procedure. By using such simulated data one may test whether RBC theory shows improvement in the experimental stage. If this procedure fails there is a strong indication the theoretical results may not translate into practical significance. If the results do show promise the next logical task is to implement the methodology to real-life data. Simulated data is considered first due to the fact that real-life data may be “dirty” and thus may be susceptible to conditioning factors that may complicate the data modelling. Such external factors may be controlled by the consideration of simulated data, providing knowledge of the true cointegration relationship present. If RBC theory translates well to simulated data and providing external factors are minimised in real-life data, the RBC
procedure is likely to be beneficial in practical situations.

Three examples are considered using simulated data. The first of these examples (Example 3) is considered because the generation involves two autoregressive processes and the possible inaccuracy of the resulting estimates when the series are not fitted properly is shown. The second of these examples (Example 4) is included because it results in the failure of the Johansen (1988) technique, while the third example (Example 5) is considered as three variables are included in the system.

The true theoretical value of the cointegration vector is first derived and the estimates of each of the cointegration vectors are compared with the true values for one simulation from each of the three examples.

Example 3

This example involves the situation where both return series, $\Delta X_t$ and $\Delta Y_t$, are generated by autoregressive models. There are no moving-average components in this generation. The objective is to compare the cointegration coefficients obtained from Johansen’s method with the RBC procedure to show they are approximately equal.

Consider two time series, $X_t$ and $Y_t$, that are generated from the following models:

$$X_t = 1.4X_{t-1} - 0.2X_{t-2} - 0.2X_{t-3} + e_t,$$

$$Y_t = 1.2Y_{t-1} - 0.2Y_{t-2} + e_t,$$

where $e_t$ is white noise with mean 0 and variance 1.
The time series may be written as

\[
(1 - B) \begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 - 0.2B^2 \end{pmatrix} \begin{pmatrix} e_t \\ e_t \end{pmatrix}
\]

The next step is to simulate data from the models in Example 3 and compare the estimates of the coefficients of the cointegrating vector using the Johansen and RBC procedures. The data is simulated using the S-Plus statistical package.

The true variance-covariance matrix of \( e_{1t} \) and \( e_{2t} \) is given by \( \Sigma_{2 \times 2} \) with all entries equal to 1. The matrix \( \Sigma \) has eigenvalues of 2 and 0. Since one eigenvalue is equal to zero - following the discussion in Lin and McCrae (1999, 2001a) - \( X_t \) and \( Y_t \) are cointegrated. The true cointegration vector of the system \( X_t \) and \( Y_t \) is given by \( (1, -2)' \) (for details on how to determine the true cointegration vector, see Lin and McCrae, 1999).

The interest focuses on the examination of whether, given a sample \( X_t \) and \( Y_t \), the RBC procedure may be implemented to estimate the cointegration vector. A sample \( (X_t, Y_t) \) of size 1000 is simulated from the models in Example 3 (once again via implementation of the S-Plus statistical package). The RBC and Johansen procedures are applied to the samples. The first task in the RBC procedure is to fit univariate ARIMA models to each of the time series \( X_t \) and \( Y_t \), both of which are \( I(1) \). The appropriate models to be fitted - using the Box-Jenkins (1976) method - are ARIMA(2,1,0) and ARIMA(1,1,0) respectively.\(^{47}\) The true value of the coefficients of the autoregressive parameters of \( X_t \) in Example 3 are 0.4 and 0.2

\(^{47}\)Obviously these are the two correct ARIMA models for \( X_t \) and \( Y_t \) from the generation of the two particular time series.
respectively, while the true value of the coefficient of the autoregressive parameter of $Y_t$ is 0.2. There are no moving-average parameters. The estimated coefficients of the autoregressive parameters of $X_t$ are 0.3442 and 0.2354. Likewise, the estimated coefficient for the autoregressive parameter of $Y_t$ is 0.1658. The estimates of the coefficients are very close to their true values.

The second task is to construct the sample covariance matrix $\hat{\Sigma}_n$, given the residual vectors $\mathbf{e}_t = (e_{1,t}, \ldots, e_{p,t})'$ and residual cross-products $e_{i,t}e_{j,t}$, $i < j \leq p$, $t = 1, 2, \ldots, n$ are both stationary. After fitting the ARIMA models to $X_t$ and $Y_t$, the residual vectors $\mathbf{e}_t = (e_{1,t}, \ldots, e_{p,t})'$ and residual cross-products $e_{i,t}e_{j,t}$, $i < j \leq p$, $t = 1, 2, \ldots, n$ are both stationary. Therefore, $\hat{\Sigma}_n$ may be implemented to estimate the true covariance matrix $\Sigma$ and is given by

$$\hat{\Sigma}_n = \begin{pmatrix} 1.0732 & 1.0739 \\ 1.0739 & 1.0756 \end{pmatrix}$$

The third task is to determine the eigenvalues of $\hat{\Sigma}_n$. The estimated eigenvalues are 2.1484 and 0.000508 respectively. Since one estimated eigenvalue is approximately equal to zero and the ratio of it to the sum of all eigenvalues is negligible (0.000236), the RBC procedure may be used to estimate the cointegration vector. The resulting eigenvectors form the matrix $A$, which is given by

$$A = \begin{pmatrix} 0.7067 & 0.7075 \\ 0.7075 & -0.7067 \end{pmatrix}$$

Furthermore, $A$ may be decomposed into $(A_1, A_2)$, where $A_1$ is formed by those
eigenvectors corresponding to the largest eigenvalue $\lambda_1 (=2.1484)$, that is,

$$A_1 = \begin{pmatrix} 0.7067 \\ 0.7075 \end{pmatrix}$$

The fourth task involves the estimation of the cointegration vector via the solving of the equation $\xi' C(1) = 0$. As mentioned previously, $C(1)$ is given by $C(B)$ with $B = 1$. In this example, $C(1)$ is equal to

$$C(1) = \begin{pmatrix} \frac{1}{0.4204} & 0 \\ 0 & \frac{1}{0.8342} \end{pmatrix} \begin{pmatrix} 0.7067 \\ 0.7075 \end{pmatrix}$$

The true cointegration vector is theoretically equal to $(1, -2)'$. The estimated cointegration vector in this simulated example is $\hat{\xi} = (1, -1.9819)'$, a very accurate estimate notwithstanding the fact that $X_t$ has estimated coefficients that are not very close to the true coefficients.

The fifth and last task in the procedure involves testing for stationarity of the linear combination of the cointegration vector. The ADF test in GIVEWIN reveals the linear combination is stationary, as does the time series plot of the linear combination, shown in Figure 5.1. The Johansen procedure - applied through the PCFIML econometric package - yields an estimate of $\hat{\xi} = (1, -2.0003)'$ in this example.

In the situation where inappropriate ARIMA models are applied to fit the time series, an inaccurate estimate of the cointegration vector will, in all likelihood, be obtained. Evidence of this may be seen in Table 5.1 where the most accurate estimate of the cointegration vector is obtained when both time series are fitted correctly.
Figure 5.1: *Time series plot of the linear combination of the cointegration vector via the RBC procedure for Example 3.*

Table 5.1: *Estimates of the cointegration vector (for four possible combinations of fitted ARIMA models) for Example 3.*

<table>
<thead>
<tr>
<th>Combination</th>
<th>$X_t$</th>
<th>$Y_t$</th>
<th>Estimate of Cointegration Vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combination 1</td>
<td>ARIMA(1,1,0)</td>
<td>ARIMA(1,1,0)</td>
<td>(1, -1.5585)</td>
</tr>
<tr>
<td>Combination 2</td>
<td>ARIMA(2,1,0)</td>
<td>ARIMA(1,1,0)</td>
<td>(1, -1.9819)</td>
</tr>
<tr>
<td>Combination 3</td>
<td>ARIMA(1,1,0)</td>
<td>ARIMA(2,1,0)</td>
<td>(1, -1.5073)</td>
</tr>
<tr>
<td>Combination 4</td>
<td>ARIMA(2,1,0)</td>
<td>ARIMA(2,1,0)</td>
<td>(1, -1.9159)</td>
</tr>
</tbody>
</table>
Example 4

This example involves the situation where both return series are generated by moving-average models. The purpose of such an example is to investigate whether the generation impacts upon the accuracy of the Johansen and RBC estimates. In this example, there are no autoregressive components in the generation of the returns.

Consider two time series, \( X_t \) and \( Y_t \), that are generated from the following models:

\[
X_t = X_{t-1} + e_t - 0.2e_{t-1},
\]
\[
Y_t = Y_{t-1} + \sqrt{2}e_t + 0.2e_{t-1},
\]

where \( e_t \) is white noise with mean 0 and variance 1. Both series \( X_t \) and \( Y_t \) are generated by a moving-average model.

The time series may be written as

\[
(1 - B) \begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} 1 - 0.2B & 0 \\ 0 & \sqrt{2} + 0.2B \end{pmatrix} \begin{pmatrix} e_t \\ e_t \end{pmatrix}
\]

The next task involves the simulation of data from the above models. The estimates of the coefficients of the cointegrating vector are then formed using the Johansen and RBC techniques. The data is simulated using the S-Plus statistical package.

The true variance-covariance matrix of \( e_{1t} \) and \( e_{2t} \) is given by \( \Sigma_{2 \times 2} \) with all entries equal to 1. The matrix \( \Sigma \) has eigenvalues of 2 and 0. Since one eigenvalue is equal to zero, \( X_t \) and \( Y_t \) are cointegrated (Lin and McCrae, 1999, 2001a). The true cointegration vector is given by \((1, -0.4956)'\).
Data is simulated from the models in Example 4 (via implementation of the S-Plus statistical package) and the RBC and Johansen procedures are applied to the simulated data of 1000 observations. In this instance, the appropriate ARIMA model to be fitted to both time series - $X_t$ and $Y_t$ - is ARIMA(0,1,1). From the fitting of this model to $X_t$ and $Y_t$, the residual vectors $e_t = (e_{1,t}, ..., e_{p,t})'$ and residual cross-products $e_{i,t}e_{j,t}, i < j \leq p, t = 1, 2, ..., n$ are both stationary. The sample covariance matrix is given by

$$
\hat{\Sigma}_n = \begin{pmatrix}
1.0743 & 1.5183 \\
1.5183 & 2.1465
\end{pmatrix}
$$

The eigenvalues of $\hat{\Sigma}_n$ are 3.2206 and 0.000234 respectively. The second estimated eigenvalue is negligible (and the ratio of it to the sum of all of the eigenvalues is also negligible) and so an estimate of the cointegration vector may be obtained by the RBC procedure. The eigenvectors form the matrix $A$, where

$$
A = \begin{pmatrix}
0.5775 & 0.8164 \\
0.8164 & -0.5775
\end{pmatrix},
$$

and where the eigenvector $(0.5775, 0.8164)'$ corresponds to the largest eigenvalue, 3.2206.

The estimated cointegration vector in this example is $\hat{\xi} = (1, -0.4941)'$, a very accurate estimate noting the true cointegration vector has been shown to be theoretically equal to $(1, -0.4956)'$. The graph of the linear combination of the cointegration vector is stationary and is shown in Figure 5.2. The ADF test also reveals stationarity of the linear combination. The traditional cointegration approach of Johansen implemented via application of the PCFIML package yields invalid estimates in this
instance. The invalid estimates arise from the singularity of the $\Sigma$ matrix (Johansen, 1988).\textsuperscript{48}

Figure 5.2: \textit{Time series plot of the linear combination of the cointegration vector via the RBC procedure for Example 4.}

---

**Example 5**

This simulation example involves the generation of three time series, $X_t$, $Y_t$, and $Z_t$. The purpose of such an example is to extend the generation to more than two variables, where the returns are generated by a mixture of autoregressive and moving-average components. The aim is to show the robustness of the RBC procedure - the RBC approach may be applied regardless of the underlying generation of the time series.

\textsuperscript{48}The resulting output is made redundant (in actual fact, the output does not reveal cointegration when clearly there is).
Consider three time series, \(X_t, Y_t\) and \(Z_t\), that are generated from the following models:

\[
X_t = 1.4X_{t-1} - 0.2X_{t-2} - 0.2X_{t-3} + e_{1,t},
\]

\[
Y_t = 1.3Y_{t-1} - 0.3Y_{t-2} + e_{1,t} + e_{2,t},
\]

\[
Z_t = 1.6Z_{t-1} - 0.6Z_{t-2} + e_{2,t} - 0.8e_{2,t-1},
\]

where \(e_{1,t}\) is white noise with mean 0 and variance 0.64 and \(e_{2,t}\) is white noise with mean 0 and variance 1. Both series \(X_t\) and \(Y_t\) are generated by an autoregressive model and the series \(Z_t\) contains both autoregressive and moving-average components.

The time series may be written as

\[
(1 - B) \begin{pmatrix} X_t \\ Y_t \\ Z_t \end{pmatrix} = \begin{pmatrix} \frac{1}{1-0.4B} & 0 & 0 \\ 0 \frac{1}{1-0.3B} & 0 \\ 0 \frac{1}{1-0.8B} \frac{1}{1-0.6B} \end{pmatrix} \begin{pmatrix} e_{1,t} \\ e_{1,t} + e_{2,t} \\ e_{2,t} \end{pmatrix}
\]

The true variance-covariance matrix is given by

\[
\Sigma = \begin{pmatrix} 0.64 & 0.64 & 0 \\ 0.64 & 1.64 & 1 \\ 0 & 1 & 1 \end{pmatrix},
\]

which has eigenvalues of 2.5173, 0.7627 and 0. Since one eigenvalue is equal to zero, \(X_t, Y_t\) and \(Z_t\) are cointegrated (Lin and McCrae, 1999, 2001a). The true cointegration vector of the system \(X_t, Y_t\) and \(Z_t\) is given by \((1, -1.75, 5)^t\).

The next step is to simulate data from the models in Example 5. Both Johansen and RBC procedures will then be applied to the data to obtain estimates of the coefficients of the cointegrating vector. The S-Plus statistical package is used to generate
the two time series and the RBC and Johansen procedures are applied to the simulated data of 1000 observations. In this instance, the appropriate ARIMA models to be fitted to \(X_t, Y_t\) and \(Z_t\) are ARIMA\((2,1,0)\), ARIMA\((1,1,0)\) and ARIMA\((1,1,1)\) respectively. After fitting the ARIMA models to \(X_t, Y_t\) and \(Z_t\), the residual vectors \(e_t = (e_{1,t}, ..., e_{p,t})'\) and residual cross-products \(e_{i,t} e_{j,t}, \; i < j \leq p, \; t = 1, 2, ..., n\) are stationary. The sample covariance matrix is given by

\[
\hat{\Sigma}_n = \begin{pmatrix}
0.6869 & 0.6919 & 0.0062 \\
0.6919 & 1.6473 & 0.9516 \\
0.0062 & 0.9516 & 0.9458
\end{pmatrix}
\]

The eigenvalues of \(\hat{\Sigma}_n\) are 2.4980, 0.7802 and 0.001755 respectively. The last estimated eigenvalue is negligible (and the ratio of it to the sum of all of the eigenvalues is also negligible). Therefore, an estimate of the cointegration vector may be obtained by the RBC procedure. The eigenvectors form the matrix \(A\), where

\[
A = \begin{pmatrix}
0.3110 & -0.7549 & 0.5774 \\
0.8097 & -0.1076 & -0.5769 \\
0.4976 & 0.6469 & 0.5778
\end{pmatrix},
\]

and where the eigenvector \((0.3110, 0.8097, 0.4976)'\) corresponds to the largest eigenvalue (2.4980) and the eigenvector \((-0.7549, -0.1076, 0.6469)'\) corresponds to the second largest eigenvalue, 0.7802. The first two columns of \(A\) are used to compose \(C(1)\).

The estimated cointegration vector is \(\hat{\xi} = (1, -1.6479, 4.8636)'\). The graph of the linear combination of the cointegration vector is stationary and is shown in Figure 5.3. The ADF test also reveals stationarity of the linear combination. The
The cointegration approach of Johansen implemented via application of the PCFIML package yields an estimate of the cointegration vector of \( \hat{\xi} = (1, -1.7664, 5.0294)' \). The resultant linear combination of cointegration vector is also stationary in this instance (both by graphical procedures and the ADF test).

The purpose of this section was to determine whether the RBC technique is applicable in practical situations. Examples 3-5 show that the RBC procedure is theoretically valid. The estimates of the cointegration vector using the RBC method are accurate irrespective of the underlying generation of the time series.

Figure 5.3: *Time series plot of the linear combination of the cointegration vector via the RBC procedure for Example 5.*
5.6 Simulated Examples

The purpose of this section is to investigate whether the cointegration vector estimates from the RBC approach vary substantially between simulations by conducting ten thousand independent simulations of the model in Example 3. The RBC procedure described previously is implemented to each simulated data set to obtain an estimate of the cointegration vector. The first component of the cointegration vector \( \xi (\xi_1) \) in all cases is fixed (equal to one) and only the second component of \( \xi (\xi_2) \) is estimated. According to the procedure outlined in Section 5.4, the negligibility of the smallest estimated eigenvalue is of interest. Therefore, for each simulation, the proportion contribution of the smallest eigenvalue (to the sum of eigenvalues) is calculated (see Examples 3-5 above). The estimated eigenvalues are not identical between replications due to the residual covariances of \( X_t \) and \( Y_t \) (and \( Z_t \) in Example 5) not being equal.

The RBC procedure is applied to ten thousand independent simulations to determine whether the resulting estimates are accurate. These simulations provide useful information in regards to the robustness (or otherwise) of the procedure. The ten thousand independent samples - \( X_t, Y_t \) - are generated from the ARIMA models described in Example 3. The correlation between the magnitude of the contribution of the smallest eigenvalue and the most accurate estimates may be evidenced by the cross-plot of the estimate of the cointegration vector via the RBC procedure and the proportion contribution of the smallest eigenvalue for 10000 simulations (see Figure 5.4).
The smaller the proportion contribution of the smallest eigenvalue, the more accurate the resulting estimate of the cointegration vector. It appears as though the resulting estimates of the cointegration vector are symmetrically dispersed around the central true value of -2. Figure 5.4 appears to be bimodal in the sense that, when the contribution of the smallest eigenvalue (relative to the sum of all eigenvalues) is relatively large, two possible events may occur - either the resulting estimate of the cointegration vector might over-estimate or the estimate may under-estimate the true cointegration vector.

Figure 5.4: Cross-plot of the estimate of the cointegration vector via the RBC procedure and the proportion contribution of the smallest eigenvalue for Example 3.
5.7 Application of the RBC Procedure to Financial Data

This section applies the RBC procedure to real-life data. The purpose is to show whether, in practical situations, the method of RBC may be invoked as an alternative to the Johansen procedure in providing accurate estimates of the coefficients of the cointegration vector. The aim is to see whether the improvement in the alternative technique that is gained with “clear” (generated) series may be translated to real-life data.

The purpose of this section in terms of the overall analysis is to determine whether the RBC procedure is a valid alternative to the Johansen technique. The RBC approach was shown to be a valid procedure in estimating the coefficients of the cointegration vector in Examples 3-5. This part of the analysis generates 1000 simulations and investigates whether the RBC technique is accurate over these simulations.

Both examples involve foreign exchange data. Example 6 tests for cointegration between spot and forward prices. Example 7 tests for cointegration between three (spot) exchange rate series in the Asian region.

Example 6

This example relates to an analysis of potential long-run equilibrium between daily spot and one week forward rates for the US dollar (relative to the UK pound) exchange rate. In Examples 3-5, the RBC procedure was shown to be a valid alternative to the Johansen technique for simulated data. In Example 4, the RBC
procedure is in fact superior to the Johansen method as the latter assumes a finite order VAR model with white noise is an adequate proxy to a finite moving-average process. In certain instances, moving-average processes may not be appropriately modelled by VAR specifications - in these cases the RBC procedure is superior to the Johansen method in terms of accuracy. The purpose of this example is to examine whether the RBC procedure is effective in empirical situations.

The period examined is between October 27, 1997 and December 22, 1999 inclusively, which provides 563 observations. Denoting $X_t$ to be the natural logarithm of the spot rates of the US Dollar (with respect to the British Pound) and $Y_t$ as the natural logarithm of the one week forward rates of the US Dollar (with respect to the British Pound), the ARIMA$(3,1,0)$ or ARIMA$(0,1,3)$ time series models seem applicable models to fit $X_t$ and $Y_t$. The fitting of these ARIMA models to $X_t$ and $Y_t$ results in the residual vectors $e_t = (e_{1,t}, ..., e_{p,t})'$ and residual cross-products $e_{i,t}e_{j,t}$, $i < j \leq p$, $t = 1, 2, ..., n$ both being stationary. Therefore, the residual vectors may be exploited to obtain the sample covariance matrix for $e_t$.

All four combinations of models are fitted to the series $X_t$ and $Y_t$ to determine which combination provides the most accurate estimate of the cointegration vector. The determination of the most accurate estimate of the cointegration vector is provided via both an ADF test and a time series plot of the resulting linear combination of the cointegration vector. Both series are non-stationary in levels and stationary in first differences.

The time series graphs of $X_t$ and $Y_t$ are shown in Figures 5.5 and 5.6 respectively. The estimates of the cointegration vector for each combination are shown in Table
5.2. The third estimate is the most accurate. The graph of the linear combination of the cointegration vector is shown in Figure 5.8. When an ADF test is applied, the linear combination of the cointegration vector for the third potential cointegration vector is the only one that reveals stationarity. The other linear combinations are not stationary. Therefore, the estimate of the cointegration vector is \((1, -1.016763)\)'.

The cointegration vector when the Johansen method is applied is \((1, -1.0118)\)'. The PCFIML output of the cointegration analysis is reported in Table 5.2. The graph of the linear combination of the cointegration vector from the Johansen procedure is shown in Figure 5.7. The graph of the linear combination of the cointegration vector obtained via the RBC technique is shown in Figure 5.8. An ADF test for stationarity

---

49The estimates in Table 5.2 arise from both time series being fitted using ARIMA models without the constant term.
on the linear combination shows significance at the 1% level of significance, thus inferring stationarity. Analysis on the linear combination of the cointegration vector from the Johansen procedure also reveals stationarity at the 1% level of significance.

**Example 7**

The method of RBC is applied to foreign exchange rates in the Asian region to examine whether the RBC technique is applicable to empirical data. In this example, there are three series of (spot) exchange rates. The first is the Malaysian Ringgit, the second is the Philippine Peso and the third series is the Thai Baht. All series are expressed in terms of the US Dollar. The time period examined commences on January 1, 1985 and terminates on December 30, 1994. There are 2610 observations (ten years data) for each series. The natural logarithms of the exchange rates are
Table 5.2: *Estimates of the cointegration vector (for four possible combinations of fitted ARIMA models) for Example 6.*

<table>
<thead>
<tr>
<th>Combination</th>
<th>$X_t$</th>
<th>$Y_t$</th>
<th>Estimate of Cointegration Vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combination 1</td>
<td>ARIMA(3,1,0)</td>
<td>ARIMA(3,1,0)</td>
<td>(1, -0.9991)</td>
</tr>
<tr>
<td>Combination 2</td>
<td>ARIMA(0,1,3)</td>
<td>ARIMA(0,1,3)</td>
<td>(1, -0.9993)</td>
</tr>
<tr>
<td>Combination 3</td>
<td>ARIMA(3,1,0)</td>
<td>ARIMA(0,1,3)</td>
<td>(1, -1.0168)</td>
</tr>
<tr>
<td>Combination 4</td>
<td>ARIMA(0,1,3)</td>
<td>ARIMA(3,1,0)</td>
<td>(1, -0.9819)</td>
</tr>
</tbody>
</table>

Figure 5.7: *Time series plot of the linear combination of the cointegration vector via the Johansen Method for Example 6.*
analysed and the time series graphs of each of the Ringatt, Peso and Baht are shown in Figures 5.9-5.11.

The most appropriate univariate Box-Jenkins ARIMA model for the Malaysian Ringgit is the ARIMA(4,1,0) model. The most appropriate Box-Jenkins ARIMA time series models for the Philippine Peso and the Thai Baht are ARIMA(1,1,2) and ARIMA(0,1,4) respectively. The resulting residual vectors $e_t = (e_{1,t}, ..., e_{p,t})'$ and residual cross-products $e_{i,t}e_{j,t}$, $i < j \leq p$, $t = 1, 2, ..., n$ are both stationary. Therefore, the residual vectors may be exploited to obtain the sample covariance matrix for $e_t$.

The resultant eigenvalues obtained from the fitting of the three ARIMA models specified above are 0.00002498, 0.00001171 and 0.00000364 respectively. These
Figure 5.9: *Time series plot of the natural logarithms of the Malaysian Ringgit for Example 7.*

![Malaysian Ringgit vs. the United States Dollar (logarithms)](image)

Figure 5.10: *Time series plot of the natural logarithms of the Philippine Peso for Example 7.*

![Philippine Peso vs. the United States Dollar (logarithms)](image)
eigenvalues indicate either zero or one cointegration vector since the second smallest 
eigenvalue is much larger than the smallest eigenvalue and not much smaller than the 
largest eigenvalue. The RBC cointegration vector estimate is $(1, 3.9640, -23.3427)'$.
The linear cointegration vector is found to be stationary using the ADF test at the 
1% level of significance. The Johansen technique gives the same conclusion. The 
estimate of the cointegration vector is given by $(1, 28.229, 298.19)'$. The linear com-
bination of cointegration vector obtained via the Johansen and RBC techniques are 
plotted in Figures 5.12 and 5.13 respectively. The cointegration vector given by the 
Johansen method assigns heavy weight on the Thai Baht, this currency plays a ma-
jor role in the system as defined by the Johansen method and the impact provided 
by the Malaysian Ringgit and Philippine Peso on the system becomes insignificant.
From this point of view, the cointegration system determined by RBC appears more appropriate.

Figure 5.12: *Time series plot of the linear combination of the Johansen cointegration vector for Example 7.*

The results in this section have confirmed the RBC technique is indeed a valid alternative to the Johansen procedure. The RBC method provides accurate estimates of the linear combination of cointegration vector, providing each individual time series is well-modelled.

### 5.8 Conclusion

The purpose of this chapter was to lay the analytical foundations for the GARCH-X model, by introducing the concept of cointegration. This chapter compared the
traditional cointegration approaches of Engle-Granger (1987) and Johansen (1988). Both of these traditional approaches restrict the time series to be well-specified by a finite order VAR model - this restriction may be particularly severe in the case of “chaotic” financial time series that appear random but actually have some deterministic elements, or when the order of the moving-average component is greater than zero and may not be adequately proxied by a finite order pure autoregressive model.

An alternative approach to identifying cointegration vectors that overcomes these limitations, under certain conditions, has been developed by Lin and McCrae (1999, 2001a). The procedure may be easily applied using widely-available standard statistical packages. The “residual-based cointegration” (RBC) approach relies on the
relationship between cointegration vectors and residual processes obtained from the fitting of univariate Box-Jenkins ARIMA models. Lin and McCrae (1999, 2001a) show that, in theory, the number and identity of the linearly independent cointegration vectors may be determined via the rank and content of the residual covariance matrix respectively. The technique is relatively simple to implement and is able to be executed using commonly available statistical packages.

This chapter showed the suitability of the RBC procedure as a technique for constructing the GARCH-X model through both simulated and real-life examples. The simulations performed in this chapter show the more negligible the contribution of the smallest eigenvalue, the more likely the estimate of the cointegration vector is closer to the true theoretical value. Given a multivariate time series and estimated eigenvalues, how may it be possible to ascertain when in fact the smaller estimated eigenvalue is negligible with respect to the larger estimated eigenvalue(s), so that cointegration is a distinct possibility? A simple plot of the time series of the linear combination of the cointegration vector(s) may well not be enough.

An associated issue for examination in future studies is the development of criteria for determining whether the smallest eigenvalue provides negligible contribution with respect to the larger eigenvalue(s). Lin and McCrae (2001b) provide a relative comparison of the variances of the non-stationary and stationary components respectively and whether the non-stationary component is negligible determines whether the system is cointegrated.

By taking into account the relationship between cointegration and univariate ARIMA models, the number of cointegration vectors may be determined via the
rank of the covariance matrix of the residual processes of univariate ARIMA models. The more negligible the contribution of the smallest estimated eigenvalue, the more accurate the resulting estimates of the cointegration vector. The RBC approach has been shown to, in certain instances, be an alternative to the Johansen procedure. The RBC method may be beneficial when the underlying time series may not be modelled appropriately by a finite order autoregression model but rather a finite order moving-average model.
Chapter 6

Modelling Conditional Moments via the GARCH-X Model

6.1 Introduction

The previous chapter discussed the implications arising from the emergence of cointegration as a powerful technique for investigating common trends in multivariate time series. Cointegration provides a precise, effective test for a fundamental property of much of finance theory - equilibrium - by enabling an investigation into the long-run as well as short-run behaviour in price series relationships, such as between spot and futures prices.

Without the added information that cointegration provides, little may be concluded about the relationship between financial time series other than to note there is substantial change in the difference between spot and futures prices in the short-run. However, with cointegration, the error correction model may be applied to
obtain information about the correction path implied for short-run changes in basis risk.

Once it can be established that two or more series are cointegrated, their dynamic structure can be exploited for further investigation. Engle and Granger (1987) show that cointegration implies an error correction representation of the component series. The following task involves analysis into the short-run situation over time and investigates the size of short-run conditional covariance changes between two series in cases where the relationship between spot and futures prices is dynamic in nature.

This chapter explains the steps in relation to forming the GARCH-X model as well as applying this model to both simulated and empirical data. Section 6.2 discusses the modelling of the conditional variance via the method of cointegration in conjunction with the GARCH specification (the GARCH-X model), and the application of this model in a hedging framework to two examples, one simulated and one empirical. The purpose of this section is to examine whether there is evidence of the effectiveness of such a model when compared to other specifications, such as the GARCH model. Section 6.3 discusses the effectiveness of the GARCH-X model both in a modelling as well as a hedging framework. Section 6.4 provides a distillation on the applicability of the analysis to future research and, in particular, to cross-hedging situations.
6.2 Applications of the GARCH-X Procedure to Effective Hedging

Lee (1994) noted that since most asset pricing theories specify conditional means as a function of conditional second moments, the consideration of the converse specification that allows for the examination of the potential relationship between disequilibrium and uncertainty in the cointegrated system may be of interest. Fama and French (1987) and Viswanath (1993) show that the basis is a significant relevant information variable. Lee (1994, pp. 375-376) states that since the error correction term may influence the conditional mean, it may also influence the conditional variance and “if disequilibrium (measured by the error correction term) is responsible for uncertainty (measured by the conditional variance) the conditional heteroscedasticity may be modelled with a function of several lagged error correction terms” to see whether some variables for the conditional means affect conditional variances. Lee (1994) considered a system of error correction models for the conditional mean and an extended bivariate GARCH model with the error correction term for the conditional variance. The model seems appropriate for testing for causality in variance through the error correction term.

The purpose of this section is to discuss the logic behind the inclusion of cointegration information (via the error correction term) in a GARCH specification. Consider a spot security and a futures contract traded on the basis of the spot security. Let $S_t$ and $F_t$ denote the natural logarithms of spot and futures prices respectively at time $t$. Since the futures contract is priced off the spot security, the
error correction term is given by

\[ z_t = S_t - \beta F_t \]

The term \( z_t \) imposes the long-run cointegration relationship between spot and futures prices and measures how the dependent variable adjusts to the previous period’s deviation from long-run equilibrium. The constant \( \beta \) is known as the cointegration parameter that links spot and futures prices (logarithms) such that the error correction term is stationary. At any given time, \( z_t \) is expected to differ from its long-run equilibrium level, known as a “disequilibrium” state. The expectation of \( z_t \) gives the long-run equilibrium relationship between \( S_t \) and \( F_t \) and short-term periods of disequilibrium occur as the observed value of \( z \) varies around its expected value. Therefore, cointegration information relating to the series \((S_t, F_t)\) may indeed be significant in modelling the conditional variances and covariances of financial asset returns.

Two examples are provided to test the applicability of the GARCH-X method in practical situations. The purpose of these examples is to outline the methodology involved in applying the GARCH-X technique and showing such a model may lead to increased hedging effectiveness. In particular, the second example tests whether the GARCH-X model increases hedging effectiveness above and beyond that obtained via the implementation of constant hedge ratios and the hedge ratios obtained via the GARCH method. The following four tasks are involved in such a process:

1. testing for skewness and leptokurtosis in spot and futures returns and any possible autocorrelation in these returns, determining the significance of any
deviation from normality.

2. testing whether cointegration exists between the two price series. If there is no such cointegration there is no logic for implementation of a model that accounts for cointegration and, as a consequence, it appears illogical to hedge using this particular futures instrument, as the two series do not appear to track each other.

3. choosing the specification of the conditional mean and conditional second moments based on information obtained from the previous two tasks.

4. constructing the GARCH-X hedge ratios and applying these hedge ratios to spot and futures returns in order to obtain values for the effectiveness of various hedging techniques.

The GARCH-X specification is used to model the conditional variances and conditional covariance for both a simulated and an empirical example. Example 8 involves data simulated from a GARCH-X model using the S-Plus statistical package. The benefits of including a simulated example is to ensure the conditional variance specification is correct and to reduce any potential errors that may stem from the inaccurate estimates of the coefficients in the conditional variance model. Example 9 uses actual futures prices for both heating oil and crude oil (as in Example 2), where it is expected the spread between the two rates is constrained (cointegrated) in the long-run, yet volatile enough for time-varying procedures to be applicable in the short-run. In this empirical example, the cointegration vector is calculated for the within-sample period and is not updated throughout the out-of-sample period.
6.2.1 Application to Simulated Data

Example 8

In this simulated example, two time series, $\Delta S_t$ and $\Delta F_t$, are generated from the GARCH-X model. The GARCH-X methodology is then applied to these return series to determine whether the application of this model results in increased hedging effectiveness when compared to static hedging procedures. The first step is to construct the model. The objective is to generate two series that are cointegrated, with both return series exhibiting GARCH effects. The series $z_t$ and $S_t$ are initially formed, the latter series generated from the GARCH-X model. The series $F_t$ is then calculated as a linear combination of $z_t$ and $S_t$. The first differences of $S_t$ (that is, the spot return series) are denoted by $e_{1,t}$. The series $e_{1,t}$ is required to be a random walk and independent of $z_t$, satisfying the condition that

$$E_{t-1}(e^2_{1,t}) = c + a_1 e^2_{t-1} + b_1 H_{11,t-1} + d_1 z^2_{t-1},$$

where $E_{t-1}$ denotes the conditional expectation given information available up to and including time $t - 1$, and $c, a_1, b_1$ and $d_1$ are (constant) coefficients. Given $e_{1,t}$ and $z_t$, assume spot and futures prices, $S_t$ and $F_t$, are generated from the following equations:

$$S_t = S_{t-1} + e_{1,t}, \quad \text{and}$$

$$z_t = S_t - \beta F_t,$$

where $z_t = \delta_t + \alpha \delta_{t-1}$ is a stationary series (when $|\alpha| < 1$) and $\delta_t$ is a sequence of normally distributed random variables. The next step is to obtain expectations of the conditional variances and conditional covariance.
Therefore,

\[ \Delta S_t = e_{1,t}, \]

\[ \Delta F_t = \frac{1}{\beta}(z_{t-1} + e_{1,t} - z_t) \]

\[ = e_{2,t}, \quad \text{and} \]

\[ H_t = \begin{pmatrix} H_{11,t} & H_{12,t} \\ H_{21,t} & H_{22,t} \end{pmatrix} \]

\[ = E_{t-1}\begin{pmatrix} e_{1,t} \\ e_{2,t} \end{pmatrix} \begin{pmatrix} e_{1,t} & e_{2,t} \end{pmatrix} \]

\[ = E_{t-1} \begin{pmatrix} e_{1,t}^2 & e_{1,t}e_{2,t} \\ e_{1,t}e_{2,t} & e_{2,t}^2 \end{pmatrix} \]

\[ = \begin{pmatrix} E_{t-1}(e_{1,t}^2) & E_{t-1}(e_{1,t}e_{2,t}) \\ E_{t-1}(e_{1,t}e_{2,t}) & E_{t-1}(e_{2,t}^2) \end{pmatrix}, \]

and

\[ E_{t-1}(e_{1,t}^2) = H_{11,t} \]

\[ E_{t-1}(e_{1,t}e_{2,t}) = \frac{H_{11,t}}{\beta} \]

\[ E_{t-1}(e_{2,t}^2) = \frac{1}{\beta^2} \left\{ [(1-\alpha)\delta_{t-1} + \alpha\delta_{t-2}]^2 + H_{11,t} + \sigma_\delta^2 \right\} \]

(see Appendix A for the derivation of these specifications).

From the above construction, \( \Delta S_t \) and \( \Delta F_t \) follow a GARCH(1,1)-X model. The series \( S_t \) and \( F_t \) are both \( I(1) \) and cointegrated (under the assumption that \( e_{1,t} \), \( e_{2,t} \) and \( z_t \) are all stationary). The conditional second moments of \( e_{1,t} \) and \( e_{2,t} \) and the conditional covariance of \( e_{1,t}e_{2,t} \) are all dependent upon the specification of the conditional second moment, \( H_{11,t} \).
The next task in examining the applicability of the GARCH-X model involves the simulation of data from the bivariate GARCH-X model outlined above. The sample size is 1210 with the first 10 observations subsequently discarded. The last 200 observations make up the out-of-sample (forecasting) period. The cointegration parameter ($\beta$) that links the two series is equal to 1, that is, the relationship $S_t - F_t = z_t$ forms a stationary series. After generating the $I(1)$ series $S_t$, $F_t$ is formed by ensuring $z_t$ is a stationary series and implementing the relationship $F_t = S_t - z_t$ to generate the sequence of futures prices. In this example, $z_t = \delta_t + 0.6\delta_{t-1}$, where $\delta_t$ is a sequence of independently and identically distributed random variables from the normal distribution with mean 0 and variance 0.5. The variance function of $\Delta S_t$ is of the form $H_t = 0.2e_{t-1}^2 + 0.8H_{t-1} + 0.1z_{t-1}^2$. The time series plots of $S_t$ and $F_t$ are shown in Figures 6.1 and 6.2 respectively while the plots of the returns, $\Delta S_t$ and $\Delta F_t$, appear in Figures 6.3 and 6.4 respectively. The two series, $S_t$ and $F_t$, are verified to be cointegrated with $\beta = 1$.

From Figures 6.3 and 6.4 it may be seen that both return series are time-varying and conditional variance specifications may be applied to model the conditional volatility. The normality assumption is tested by examining the measures of skewness and kurtosis. There exists significant skewness (at the 1% level of significance) and kurtosis (at the 1% level of significance) in both $\Delta S_t$ and $\Delta F_t$. Both return series are also not normally distributed (at the 1% level of significance) and exhibit significant autocorrelation (once again at the 1% level of significance). The absolute values of $\Delta S_t$, $\Delta F_t$, $\Delta S_t^2$ and $\Delta F_t^2$ exhibit similar autocorrelation (at the 1% level of significance). The sign for the skewness of both returns series is negative, indicating
Figure 6.1: Time series plot of the spot prices for Example 8.

Figure 6.2: Time series plot of the futures prices for Example 8.
Figure 6.3: *Time series plot of the spot returns for Example 8.*

![Time series plot of spot returns](image)

Figure 6.4: *Time series plot of the futures returns for Example 8.*

![Time series plot of futures returns](image)
the distribution of each return series is skewed left. Likewise, both $\Delta S_t$ and $\Delta F_t$ exhibit significant kurtosis, indicating the distributions of both spot and futures returns are heavy-tailed. The Jarque-Bera test for normality also reveals significant non-normality in $\Delta S_t$ and $\Delta F_t$ (at the 1% level of significance). The Ljung-Box and Lagrange-Multiplier tests reveal $\Delta S_t$ and $\Delta F_t$ are both autocorrelated (at the 1% level of significance) and therefore significantly impacted by ARCH effects. The summary statistics are included in Table 6.1.

Table 6.1: Summary statistics for the spot and futures returns for Example 8.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>$\Delta S_t$</th>
<th>$\Delta F_t$</th>
<th>5% Critical Values</th>
<th>1% Critical Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>1000</td>
<td>1000</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.802$^a$</td>
<td>-0.834$^a$</td>
<td>(-0.152,0.152)</td>
<td>(-0.200,0.200)</td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>8996.124$^a$</td>
<td>8508.730$^a$</td>
<td>5.991</td>
<td>9.210</td>
</tr>
<tr>
<td>Ljung-Box $Q(24)$</td>
<td>329.428$^a$</td>
<td>317.495$^a$</td>
<td>36.415</td>
<td>42.980</td>
</tr>
<tr>
<td>Ljung-Box $</td>
<td>Q(24)</td>
<td>$</td>
<td>9759.903$^a$</td>
<td>9629.152$^a$</td>
</tr>
<tr>
<td>Ljung-Box $Q^2(24)$</td>
<td>4592.962$^a$</td>
<td>4606.894$^a$</td>
<td>36.415</td>
<td>42.980</td>
</tr>
<tr>
<td>Lagrange-Multiplier</td>
<td>503.901$^a$</td>
<td>504.084$^a$</td>
<td>36.415</td>
<td>42.980</td>
</tr>
</tbody>
</table>

Notes:
1. $^a$ denotes significance at the 1% level of significance.
2. $^b$ denotes significance at the 5% level of significance.

The final task in this example involves the application of various hedging techniques to the returns series. The purpose of these applications is to enable a comparison of the hedge ratios generated by these constant ratio techniques with that of a GARCH-X model in order to compare relative hedging effectiveness. The performances of various hedging methods are compared as in Chapter 4. Daily returns
are constructed as implied by the computed hedge ratios and the variance of the returns of the constructed portfolios are calculated over the entire sample period. It was shown in Chapter 4 that when the hedge ratios are unstable, allowance for such stochastic movements substantially increases hedging effectiveness by reducing the volatility of the hedged portfolio. The conditional variance of any such hedge at time \( t \) is equal to

\[
H_{11,t} - 2hH_{12,t} + h^2H_{22,t},
\]

where \( h \) denotes the hedge ratio. In this analysis, \( h \) may either be equal to the minimum-variance hedge ratio, \( h_{MV} \), the naive hedge ratio, 1, or the dynamic hedge ratio, \( h_t \). In Chapter 4 the minimum-variance hedging technique was shown to provide a better hedge than the naive procedure when

\[
(h_{MV} + 1)H_{22,t} - 2H_{12,t} > 0
\]

That is, when the quantity \((h_{MV} + 1)H_{22,t} > 2H_{12,t}\), the minimum-variance hedge is preferred over the naive procedure (in terms of greater risk-reduction). Similarly, the dynamic hedge produces a more effective hedge than the naive technique when

\[
-(H_{12,t} - H_{22,t})^2 < 0
\]

The dynamic hedge produces a better hedge than the naive hedge as the conditional covariance between the spot and futures returns deviates from the conditional variance of the futures returns. The naive hedge produces an adequate hedge when the quantities \( H_{12,t} \) and \( H_{22,t} \) are approximately equal (that is, near-perfect correlation between spot and futures returns). Finally, the dynamic technique provides a
better hedge than the minimum-variance technique when

\[-\left(\frac{H_{12,t}}{H_{22,t}} - h_{MV}\right)^2 < 0\]

In Example 8, the true form of the variance function is known. Therefore, the conditional variance of each of the hedging methods (naive, minimum-variance and GARCH-X) may be calculated and compared to determine the effectiveness of each method. If the error correction term is influential in modelling the conditional covariance and conditional variances, a GARCH model without cointegration will result in model mis-specification.

The focus turns to the effectiveness of each type of hedge. In Example 8, the true type of variance function is known as well as the resulting coefficients in this function. The time series plot of the GARCH-X hedge ratios for the out-of-sample period is shown in Figure 6.5.

The GARCH-X hedge ratios vary from 0.4934 to 0.9708 in the out-of-sample period. The mean value of these hedge ratios is 0.8134. The minimum-variance hedge ratio, on the other hand, generates a ratio of 1.0009 within-sample and this ratio is held constant throughout the out-of-sample period. In this example, the minimum-variance method overestimates the correlation between spot and futures returns in most periods. The plots of the differences between the conditional variances for all combinations of hedging rules appear in Figures 6.6-6.8. In this example, the hedging effectiveness value for the GARCH-X method is 0.8122. This value compares favourably to the hedging effectiveness values obtained via both the naive and minimum-variance hedges, which are 0.7596 and 0.7591 respectively.
Figure 6.5: *Time series plot of the hedge ratios obtained via implementation of the bi-variate GARCH-X model for Example 8.*

Figure 6.6: *Conditional variance of the time-varying (GARCH-X) hedge minus the conditional variance of the minimum-variance hedge for Example 8.*
Figure 6.7: Conditional variance of the time-varying (GARCH-X) hedge minus the conditional variance of the naive hedge for Example 8.

Figure 6.8: Conditional variance of the minimum-variance hedge minus the conditional variance of the naive hedge for Example 8.
6.2.2 Application to Financial Data

Example 9

The purpose of this example is to examine whether the GARCH-X model is effective in empirical situations. The GARCH-X specification was shown to be beneficial using simulated data (Example 8) and now will be tested using financial data. This example involves the same heating oil and crude oil futures data as in Example 2. The conditional variance model, however, is extended to include a function of the error correction term. The conditional mean is fitted using Equation (4.4), while the conditional variance specification is given by Equation (3.10). Via implementation of the PCFIML econometric package, the cointegration parameter, $\beta$, that links the two series is found to be equal to 0.88466 (and is significant at the 1% level of significance). Therefore, $\ln(H) - 0.88466 * \ln(C)$ provides a stationary series, where $H$ and $C$ denote the heating oil and crude oil prices respectively and $\ln$ denotes the natural logarithm. The lag length chosen is equal to 1 and ensures there is no serial correlation in the residuals. The output is omitted here but is available on request. The time series graph of the linear combination of the cointegration vector is shown in Figure 6.9.

The diagonal representations of the GARCH and GARCH-X models are both considered as they are time-varying models and one may like to investigate whether the GARCH-X model has potential to produce a better hedge and/or provide more information in regards to the dynamic properties of the spot and futures returns when compared to its GARCH counterpart (Kavussanos and Nomikos, 2000a, 210).
Figure 6.9: *Time series plot of the linear combination of the Johansen cointegration vector for Example 9.*

2000b). The estimates of the GARCH and GARCH-X coefficients within-sample, calculated via the econometric package RATS using the BFGS algorithm, are included in Table 6.2.

The parameters $D_{11}$ and $D_{12}$ - describing the influence of the error correction term on the conditional variance - are both significant (at the 1% level of significance) when analysing the whole sample, indicating these terms have potential predictive power in modelling the conditional variance-covariance matrix of the returns. Therefore, last period’s equilibrium error has significant impact on the adjustment process of the subsequent returns. Within-sample, only the parameter $D_{12}$ is significant (at the 1% level of significance). The log-likelihood function values are higher for the GARCH-X model than the GARCH model (both within-sample and analysing the
Table 6.2: Parameter estimates resulting from the implementation of the bivariate GARCH and bivariate GARCH-X models respectively for Example 9.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Within-Sample GARCH</th>
<th>Within-Sample GARCH-X</th>
<th>Total-Sample GARCH</th>
<th>Total-Sample GARCH-X</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_1$</td>
<td>-0.0003$^a$</td>
<td>-0.0004$^a$</td>
<td>-0.0003$^a$</td>
<td>-0.0004$^a$</td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>-0.0001$^a$</td>
<td>-0.0002$^a$</td>
<td>-0.0001$^a$</td>
<td>-0.0002$^a$</td>
</tr>
<tr>
<td>$C_{11}$</td>
<td>0.0025$^a$</td>
<td>0.0021$^a$</td>
<td>0.0025$^a$</td>
<td>0.0024$^a$</td>
</tr>
<tr>
<td>$C_{12}$</td>
<td>0.0023$^a$</td>
<td>0.0012</td>
<td>0.0024$^a$</td>
<td>0.0018$^a$</td>
</tr>
<tr>
<td>$C_{22}$</td>
<td>0.0017$^a$</td>
<td>0.0000</td>
<td>0.0017$^a$</td>
<td>0.0000</td>
</tr>
<tr>
<td>$A_{11}$</td>
<td>0.2764$^a$</td>
<td>0.2735$^a$</td>
<td>0.2778$^a$</td>
<td>0.2760$^a$</td>
</tr>
<tr>
<td>$A_{22}$</td>
<td>0.2306$^a$</td>
<td>0.2251$^a$</td>
<td>0.2691$^a$</td>
<td>0.2695$^a$</td>
</tr>
<tr>
<td>$B_{11}$</td>
<td>0.9543$^a$</td>
<td>0.9550$^a$</td>
<td>0.9543$^a$</td>
<td>0.9546$^a$</td>
</tr>
<tr>
<td>$B_{22}$</td>
<td>0.9629$^a$</td>
<td>0.9628$^a$</td>
<td>0.9554$^a$</td>
<td>0.9536$^a$</td>
</tr>
<tr>
<td>$D_{11}$</td>
<td>-</td>
<td>-0.0009</td>
<td>-</td>
<td>-0.0006$^a$</td>
</tr>
<tr>
<td>$D_{12}$</td>
<td>-</td>
<td>-0.0020$^a$</td>
<td>-</td>
<td>-0.0019$^a$</td>
</tr>
<tr>
<td>Log-likelihood Function Value</td>
<td>9380.77</td>
<td>9383.70</td>
<td>9833.70</td>
<td>9835.95</td>
</tr>
</tbody>
</table>

Notes:
1. $a$ denotes significance at the 1% level of significance.
2. $b$ denotes significance at the 5% level of significance.
total sample). This indicates that the GARCH-X model is a more adequate model in analysing the conditional variance-covariance matrix of the returns.

Implementing a GARCH model without the restriction that the matrices \( \mathbf{A} \) and \( \mathbf{B} \) are diagonal achieves a similar result to the diagonal GARCH-X model, that is, both parameters \( D_{11} \) and \( D_{12} \) are significant for the whole sample and the parameter \( D_{12} \) is significant in-sample (all at the 1% level of significance). Therefore, the (square of the) error correction term remains a significant explanatory variable for the conditional variance-covariance matrix even after allowing all components of the matrices \( \mathbf{A} \) and \( \mathbf{B} \) to vary through time.

In spite of the superiority of the GARCH-X model in modelling the variance-covariance matrix of the returns in this example, the hedge ratios obtained by the GARCH-X formulation are similar to those obtained by the GARCH specification. Therefore, the benefits of incorporating cointegration information into the formation and adjustment of hedge ratios may, in some instances, be minimal.

The main conclusion reached from Examples 8 and 9 is that the GARCH-X model may be utilised in practical situations to provide greater knowledge of how the individual components in the variance-covariance matrix behave over time. However, this may not necessarily translate into increased hedging effectiveness. The GARCH-X model may be more effective in inefficient markets that have large volatilities or markets with no futures contracts, thus invoking the need for cross-hedging.
6.3 Effectiveness of the GARCH-X Model

The technique of cointegration when applied to common trends in multivariate time series allows the retention and modelling of both long-run and short-run dynamics in a system. This represents a major potential improvement over GARCH models in situations where the series may change in the short-run, even though there exists a long-run equilibrium relationship between spot and futures prices. These short-run fluctuations in the relationship between the two (cointegrated) series do not persist indefinitely. The relationship between the two prices eventually reverts back to a long-run equilibrium level.

The studies of Fama and French (1987) and Viswanath (1993) show the basis (or more generally, the error correction term) can influence the conditional mean. Lee (1994) hypothesised that since the error correction term influences the conditional mean, it may also impact on the conditional variance. The GARCH-X model provides a measure for investigating the adjustments in the short-run shocks or deviation from the long-run equilibrium level. The measure used is the square of the long-run cointegration relationship. The square is then taken as a predictor variable of the conditional variance-covariance matrix of the returns.

This chapter discussed the modelling of the conditional variance via the GARCH-X model of Lee (1994). Lee (1994) suggested that since conditional means are usually specified as a function of conditional second moments, an examination of the converse specification may distinguish between the notions of disequilibrium and uncertainty in a system of cointegrated variables. The empirical example in
this chapter indicated that whilst the GARCH-X model explains the conditional
variances of, and the conditional covariance between, spot and futures returns better
than the GARCH model, there is no substantial difference in the hedge ratios using
the GARCH and GARCH-X models.

This examination is deliberately selective as it is not possible to deal extensively
with all the issues involved in the application of GARCH-X models in a hedging
framework. There may well be a possible link between the length of the out-of-
sample time period and the effectiveness of the GARCH-X model (with respect to
its GARCH counterpart). A longer out-of-sample time period may show short-run
deviations between the two prices are corrected in later time periods. The out-of-
sample period may also be too short to be influenced by the cointegration parameter.
However, such issues must be left for later research. The major purpose of this thesis
is to lay down the methodology for deriving a GARCH-X model and briefly test its
performance in relation to other models.

6.4 Conclusion

Theoretically, cointegration can provide substantial improvement when modelling
the variance-covariance matrix of two or more returns series. The incorporation
of cointegration relationships into the conditional variance model provides added
information that enables inclusion of long-run behaviour between two or more coin-
tegrated series. The addition also allows analysis of the dynamics of short-run “price
shocks” to the system through an error correction model. Both are significant con-
tributions to the modelling of time series behaviour through GARCH models.

Several factors may mitigate against any great improvement in a hedging framework:

1. the estimates of variance and volatility may be accurate via the implementation of static techniques. There would therefore be little room for improvement and often this would typify an efficient market where the potential for a better hedge is low.

2. even where there is potential for improvement, application of the GARCH model (as in Example 9) provides this improvement, so the GARCH-X model may not outperform the GARCH model.

Cointegration brings added information about long-run (and short-run) correlations between asset prices. It also permits analysis of the dynamics of deviations from this long-run equilibrium level. By using cointegration, investors may obtain added information in forming and/or progressively re-adjusting hedges. This re-adjustment may help in maintaining or improving the hedging effectiveness since new information impacts on asset prices.

In modelling both conditional variances and conditional covariances through the GARCH-X model, hedge ratios may be determined using an alternative specification to the GARCH model in explaining the correlation between returns. As a result, substantial basis risk may have important implications in cross-hedging situations. Cointegration may help create more effective hedges than GARCH models because the impact of cointegration on the adjustment process between spot and futures
returns allows the variance of the hedged portfolio to be minimised, creating a more effective hedge. The simulated example emphasised the substantial improvement the GARCH-X model provides in a hedging framework. Applying a GARCH model in such instances would result in a mis-specification of the variance-covariance matrix of returns.

It appears as though some unmodelled conditional heteroscedasticity in the GARCH(1,1) model may be explained by a function of the spread. This has been shown to be true for an empirical example involving commodity futures data. The magnitude of deviations from the cointegration level may provide added information about the relationship between two rates. However, this does not necessarily translate to a more effective hedge. The GARCH-X model appears to explain the relationship between disequilibrium and conditional volatility better than the GARCH(1,1) specification and should be utilised as it is highly probable that such a model will not provide a less inferior hedging strategy when compared to the GARCH model alone.
Chapter 7

Conclusion

Time variation in the conditional variance of financial time series is important in a number of financial applications, for example, pricing derivatives, calculating measures of risk, and hedging against portfolio risk. Therefore, there has been enormous interest amongst researchers and practitioners in modelling the conditional variance. Traditional theoretical and empirical investigations into optimal hedging make generalised assumptions in regards to the time series characteristics of component spot and derivative prices. Such methods assume basis risk is characterised by constant variance over time, imposing the restriction of a constant joint distribution of spot and futures price changes.

The assumption of small, stable, time-invariant basis risk may lead to suboptimal hedging decisions in periods of high basis volatility. Static hedges are calculated and applied without adjustments going forward, ignoring new information that arrives at the market. Where this new information substantially alters the relationship between spot and futures returns through time, static hedges are likely
to be sub-optimal. Consequently, conditional hedging models should be considered.

The objective of this thesis was to provide knowledge that may help in ensuring that underlying models do capture as much information in regards to time series behaviour as possible. The particular issue addressed is the potential existence of conditional volatility found to be a very common characteristic among prices of financial time series. The specific issue investigated is the manner in which conditional volatility may be integrated into the fundamental time series modelling process. The purpose of the thesis is to successfully develop a method to achieve this integration of conditional volatility and to indicate when such a development is likely to be beneficial - both theoretically and practically.

The research within this dissertation investigated conditional variance models that accommodate time series characteristics of both short and long-term behaviour commonly found in spot and futures returns. GARCH models decompose the total volatility into permanent and transitory components. Incorporating dynamic behaviour into the basic underlying GARCH model allows a greater understanding of the dynamics behind the generation of spot and futures returns and the relationship between these returns over time.

In an attempt to achieve the objectives outlined above, the thesis makes several substantial contributions to the literature on conditional variance models applied in a hedging framework. The first contribution of the research involves the development of statistical criteria that enables the comparison of different constant hedge ratios. The criteria developed in this research is based on conditional variance measures - this allows a hedger to determine the most effective constant hedge ratio and to
alternate between different static hedge ratios at appropriate points in time.

The analysis and testing of statistical criteria is important in the overall context of calculating and maintaining effective hedge ratios. Dynamic hedging approaches in minimising risk are compared to various conventional static hedge ratios and, where the hedge ratios are unstable, hedging procedures that allow for such stochastic volatility substantially increase the effectiveness of the hedge.

The second contribution of this analysis to the literature involves the introduction of an alternative constant hedge ratio to the naive and minimum-variance hedges. This innovative hedge ratio is termed the “forecasted hedge”. The forecasted hedge is obtained from the forecasting curves of the conditional covariance between spot and futures returns and the conditional variance of the futures returns. The minimum-variance hedge ratio minimises basis risk in-sample. However, there is no assurance the minimum-variance hedge ratio is the most effective hedge (in terms of risk minimisation) out-of-sample.

The forecasted hedge may indeed be superior to the minimum-variance hedge as it utilises dynamic information from the GARCH model and, based on this information, obtains forecasting curves for both the variances of spot and futures returns and the covariance between these returns. However, the resulting hedge is a constant hedge and does not involve complex re-calculation and re-balancing. The forecasted hedge is of potential interest to the hedger who would like to implement a constant hedge but does not wish to be constrained to choosing either the naive or minimum-variance hedge. Alternating between the forecasted hedge, naive and minimum-variance hedges appears logical since statistical criteria have been pro-
vided that may distinguish between the effectiveness of various constant hedges at any point in time.

A third contribution this thesis makes to the literature involves the confirmation of the usefulness of the GARCH-X model. The bivariate GARCH-X specification extends the GARCH model by incorporating information from cointegration between spot and futures prices into the GARCH model. The cointegration relationship (via the error correction term) is incorporated into the modelling of the conditional variances and the conditional covariance and is found to provide added predictive capacity over the GARCH model. In the empirical example (Example 9), though the square of the error correction term is found to provide added information in regards to the variance-covariance matrix of the returns, there is subsequently little difference between the hedge ratios generated via this model (the GARCH-X model) and the GARCH model.

The GARCH-X model is potentially superior to the basic GARCH specification since the GARCH-X formulation provides information relating to both the long-run equilibrium level as well as short-run deviations to this equilibrium level. The cointegration relationship may be an important predictor variable in the modelling of the conditional second moments. Whilst theoretically there appears to be room for substantial improvement by incorporating cointegration into the modelling of the variance-covariance matrix of returns, the improvements provided via application of the GARCH-X model (with respect to the GARCH model) may possibly exist in a hedging framework in cases involving non-existent or illiquid futures markets, where time varying basis risk is more pronounced due to the imperfect connection between
the futures market used to cross-hedge the underlying asset and the underlying spot market asset itself.

The fourth major contribution of this analysis involves an alternative specification of the cointegration relationship. The cointegration relationship is usually determined using the Johansen procedure. This research shows another method for obtaining cointegration relationships - known as the residual-based cointegration technique - that is both applicable and preferable in many situations. Theoretically, the number and identity of the linearly independent cointegration vectors may be determined via the rank and content of the residual covariance matrix respectively. This alternative construction of cointegration vectors does not impose the restriction the time series must be well-specified by a finite order VAR model. In cases where the time series contain some underlying deterministic elements or where there exists a moving-average component, the residual-based cointegration technique may provide more accurate estimates of the underlying cointegration vector.

The generality of the contribution of the dissertation is important for several reasons:

1. the contribution is significant to the theoretical development of improved time series modelling by developing and implementing a method that will better capture conditional volatility than previous models.

2. while the analysis is performed in a specific market context - hedging commodity price exposure - the theoretical extension is quite general and can be applied to any hedging market, product or context where conditional volatility
is likely to be a significant characteristic of time series behaviour.

3. the contribution of this thesis also applies in other decision making contexts that rely on fundamental time series models to reflect all significant characteristics and information inherent in cointegrated time series.

4. the potential for the GARCH-X model goes beyond integrating conditional volatility - it also gives improved insight into the dynamics of short-run deviations from long-run equilibrium processes.

There are several limitations of the thesis. These questions, albeit interesting, were not intended to be answered in the dissertation. These issues may be addressed in future research in order to further the work achieved within this analysis. Such issues involve, but are not limited to, the

1. consideration of transaction costs that are incurred in forming, and progressively re-adjusting, the hedge.

2. incorporation of risk/return payoff functions that do not necessarily enforce risk minimisation.

3. application of the GARCH-X specification in cross-hedging situations, for example, in foreign exchange hedging of certain Asian currencies with illiquid/inefficient futures markets.
Bibliography


Appendices

Appendix A

The conditional covariance and conditional variances in Section 6.2.1 were given by

\[
E_{t-1}(e_{1,t}^2) = H_{11,t}
\]

\[
E_{t-1}(e_{1,t}e_{2,t}) = \frac{H_{11,t}}{\beta}
\]

\[
E_{t-1}(e_{2,t}^2) = \frac{1}{\beta^2} \{ [(1 - \alpha)\delta_{t-1} + \alpha\delta_{t-2}]^2 + H_{11,t} + \sigma_2^2 \}
\]

The above equations for the conditional covariance and conditional variances were derived as follows:

\[
E_{t-1}(e_{1,t}^2) = H_{11,t-1}
\]

\[
E_{t-1}(e_{1,t}e_{2,t}) = E_{t-1}(e_{1,t}\frac{1}{\beta}(z_{t-1} + e_{1,t} - z_t))
\]

\[
= \frac{1}{\beta} E_{t-1}[e_{1,t}z_{t-1} + e_{1,t}^2 - e_{1,t}z_t]
\]

\[
= \frac{1}{\beta} E_{t-1}[e_{1,t}(\delta_{t-1} + \alpha\delta_{t-2})] + \frac{1}{\beta} E_{t-1}(e_{1,t}^2) - \frac{1}{\beta} E_{t-1}[e_{1,t}(\delta_t + \alpha\delta_{t-1})]
\]

\[
= \frac{1}{\beta} E_{t-1}(e_{1,t}^2)
\]

\[
= \frac{H_{11,t-1}}{\beta}
\]

\[
E_{t-1}(e_{2,t}^2) = E_{t-1}\{ \frac{1}{\beta} [(\delta_{t-1} + \alpha\delta_{t-2}) + e_{1,t} - (\delta_t + \alpha\delta_{t-1})]^2 \}
\]

\[
= \frac{1}{\beta^2} E_{t-1}[(\delta_{t-1} + \alpha\delta_{t-2})^2 + e_{1,t}^2 + (\delta_t + \alpha\delta_{t-1})^2]
\]

\[- 2(\delta_t + \alpha\delta_{t-1})(\delta_{t-1} + \alpha\delta_{t-2})
\]

\[- 2(\delta_t + \alpha\delta_{t-1})e_{1,t} + 2(\delta_{t-1} + \alpha\delta_{t-2})e_{1,t}]
\]
\[
\begin{align*}
&= \frac{1}{\beta^2}[(\delta_{t-1} + \alpha \delta_{t-2})^2 + H_{11,t-1} + E_{t-1}(\delta_t^2 + 2\alpha \delta_t \delta_{t-1} + \alpha^2 \delta_{t-1}^2) \\
&- 2(\delta_{t-1} + \alpha \delta_{t-2})E_{t-1}(\delta_t + \alpha \delta_{t-1}) - 2E_{t-1}[(\delta_t + \alpha \delta_{t-1})e_{1,t}] \\
&+ 2(\delta_{t-1} + \alpha \delta_{t-2})E_{t-1}(e_{1,t})] \\
&= \frac{1}{\beta^2}[(\delta_{t-1} + \alpha \delta_{t-2})^2 + H_{11,t-1} + \sigma^2_{\delta} + \alpha^2 \delta_{t-1}^2 - 2(\delta_{t-1} + \alpha \delta_{t-2})\alpha \delta_{t-1}] \\
&= \frac{1}{\beta^2}[((\delta_{t-1} + \alpha \delta_{t-2}) - \alpha \delta_{t-1})^2 + H_{11,t-1} + \sigma^2_{\delta}] \\
&= \frac{1}{\beta^2}[((1 - \alpha)\delta_{t-1} + \alpha \delta_{t-2})^2 + H_{11,t-1} + \sigma^2_{\delta}] 
\end{align*}
\]
Appendix B

PCFIML output of the cointegration analysis (Johansen procedure) for Example 6

Note: * and ** denote significance at the 5% and 1% levels respectively.

Cointegration analysis 2 to 563

<table>
<thead>
<tr>
<th>eigenvalue</th>
<th>loglik for rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>8689.86</td>
<td>0</td>
</tr>
<tr>
<td>0.0279362</td>
<td>8697.83</td>
</tr>
<tr>
<td>0.00145882</td>
<td>8698.24</td>
</tr>
</tbody>
</table>

Ho:rank=p -Tlog(1-\mu) using T-nm 95% -T\Sum log(.) using T-nm 95%

p == 0 15.92* 15.87* 14.1 16.74* 16.68* 15.4
p <= 1 0.8205 0.8175 3.8 0.8205 0.8175 3.8

standardised \beta' eigenvectors

<table>
<thead>
<tr>
<th>LNUSspot</th>
<th>LNUSfutu</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0000</td>
<td>-1.0118</td>
</tr>
<tr>
<td>-1.0025</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

standardised \alpha coefficients

<table>
<thead>
<tr>
<th>LNUSspot</th>
<th>3.4170</th>
<th>0.41913</th>
</tr>
</thead>
<tbody>
<tr>
<td>LNUSfutu</td>
<td>3.4395</td>
<td>0.41213</td>
</tr>
</tbody>
</table>

long-run matrix Po=\alpha*\beta', rank 2

<table>
<thead>
<tr>
<th>LNUSspot</th>
<th>LNUSfutu</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.9968</td>
<td>-3.0380</td>
</tr>
<tr>
<td>3.0264</td>
<td>-3.0679</td>
</tr>
</tbody>
</table>

Number of lags used in the analysis: 1
Variables entered unrestricted:
Constant
PCFIML output of the cointegration analysis (Johansen procedure) for Example 7

Note: * and ** denote significance at the 5% and 1% levels respectively.

Cointegration analysis 2 to 2610

eigenvalue loglik for rank

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>44539.2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>44616.2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>44623.2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>44625.2</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

Ho: rank=p -Tlog(1-\mu) using T-nm 95% -T\Sum log(.) using T-nm 95%
p == 0 154** 153.8** 21.0 172.1** 171.9** 29.7
p <= 1 14.03 14.01 14.1 18.06* 18.04* 15.4
p <= 2 4.038* 4.033* 3.8 4.038* 4.033* 3.8

standardised \beta' eigenvectors

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>LNMALA</td>
<td>-3.0439</td>
<td>28.229</td>
<td>298.19</td>
</tr>
<tr>
<td>LNPHIL</td>
<td>0.13414</td>
<td>1.0000</td>
<td>1.0115</td>
</tr>
<tr>
<td>LNTTHAI</td>
<td>0.13414</td>
<td>-1.2240</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

standardised \alpha coefficients

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>LNMALA</td>
<td>-0.00027493</td>
<td>0.0032365</td>
<td>0.00054010</td>
</tr>
<tr>
<td>LNPHIL</td>
<td>-0.00038269</td>
<td>-0.00027406</td>
<td>0.0018393</td>
</tr>
<tr>
<td>LNTTHAI</td>
<td>-0.00033567-1.7675e-006</td>
<td>-0.00046439</td>
<td></td>
</tr>
</tbody>
</table>

long-run matrix Po=\alpha*\beta', rank 3

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>LNMALA</td>
<td>-0.010054</td>
<td>-0.0051856</td>
<td>-0.078167</td>
</tr>
<tr>
<td>LNPHIL</td>
<td>0.00069821</td>
<td>-0.013328</td>
<td>-0.11255</td>
</tr>
<tr>
<td>LNTTHAI</td>
<td>-0.00039258</td>
<td>-0.0089088</td>
<td>-0.10056</td>
</tr>
</tbody>
</table>

Number of lags used in the analysis: 1 Variables entered unrestricted:

Constant
PCFIML output of the cointegration analysis (Johansen procedure) for Example 8

Note: * and ** denote significance at the 5% and 1% levels respectively.

Cointegration analysis 2 to 1000

<table>
<thead>
<tr>
<th>Eigenvalue</th>
<th>Loglik for rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.495992</td>
<td>-1630.63</td>
</tr>
<tr>
<td>0.00802488</td>
<td>-1626.60</td>
</tr>
</tbody>
</table>

Ho: rank=p  

\(-T\log(1-\mu)\) using T-nm 95% \(-T\sum \log(.)\) using T-nm 95%

<table>
<thead>
<tr>
<th>p == 0</th>
<th>684.5**</th>
<th>683.1**</th>
<th>14.1</th>
<th>692.5**</th>
<th>691.1**</th>
<th>15.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>p &lt;= 1</td>
<td>8.049**</td>
<td>8.033**</td>
<td>3.8</td>
<td>8.049**</td>
<td>8.033**</td>
<td>3.8</td>
</tr>
</tbody>
</table>

Standardised \(\beta\)' eigenvectors

<table>
<thead>
<tr>
<th>SPOT</th>
<th>FUTURES</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0000</td>
<td>-1.0006</td>
</tr>
<tr>
<td>4.5068</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Standardised \(\alpha\) coefficients

<table>
<thead>
<tr>
<th>SPOT</th>
<th>FUTURES</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.11527</td>
<td>-0.0029015</td>
</tr>
<tr>
<td>1.1005</td>
<td>-0.0028208</td>
</tr>
</tbody>
</table>

Long-run matrix \(P_0=\alpha*\beta\)', rank 2

<table>
<thead>
<tr>
<th>SPOT</th>
<th>FUTURES</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10219</td>
<td>-0.11824</td>
</tr>
<tr>
<td>1.0878</td>
<td>-1.1040</td>
</tr>
</tbody>
</table>

Number of lags used in the analysis: 1 Variables entered unrestricted:

Constant
PCFIML output of the cointegration analysis (Johansen procedure) for Example 9

Note: * and ** denote significance at the 5% and 1% levels respectively.

Cointegration analysis 2 to 1252

<table>
<thead>
<tr>
<th>eigenvalue</th>
<th>loglik for rank</th>
<th>eigenvalue</th>
<th>loglik for rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0158427</td>
<td>10314.6 0</td>
<td>0.0049735</td>
<td>10324.6 1</td>
</tr>
</tbody>
</table>

Ho: rank = p -T log(1-\mu) using T-nm 95% -T\sum log(\cdot) using T-nm 95%
p == 0 19.98** 19.95** 14.1 26.22** 26.17** 15.4

standardised \( \beta' \) eigenvectors

<table>
<thead>
<tr>
<th>LNHEAT</th>
<th>LNCRUDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0000</td>
<td>-0.88466</td>
</tr>
<tr>
<td>0.63982</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

standardised \( \alpha \) coefficients

<table>
<thead>
<tr>
<th>LNHEAT</th>
<th>LNCRUDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.020826</td>
<td>-0.0056688</td>
</tr>
<tr>
<td>0.012073</td>
<td>-0.0065155</td>
</tr>
</tbody>
</table>

long-run matrix Po=\( \alpha*\beta' \), rank 2

<table>
<thead>
<tr>
<th>LNHEAT</th>
<th>LNCRUDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.024453</td>
<td>0.012755</td>
</tr>
<tr>
<td>0.0079042</td>
<td>-0.017196</td>
</tr>
</tbody>
</table>

Number of lags used in the analysis: 1

Variables entered unrestricted:

Constant
Appendix C

S-Plus program implemented in Example 1:

e1_vector()
e2_vector()
e1f_vector()
e2f_vector()
H11_vector()
H12_vector()
H22_vector()
hrt_vector()
cvtv_vector()
cvmv_vector()
cvna_vector()
cvfo_vector()
cvco_vector()

var_matrix()
varini_matrix()
H1_matrix()
ABCCBA1_matrix()
ABCCBA2_matrix()
ABCCBA3_matrix()
ABCCBA1f_matrix()
ABCCBA2f_matrix()
ABCCBA3f_matrix()
Hfore_matrix()

e_data.frame()

set.seed(1)

n_1000
f_100

A_matrix(c(0.4,0,0,0.4),2,2)
B_matrix(c(0.9,0,0,0.9),2,2)
C_matrix(c(0.02,0.02,0,0.025),2,2)
model.bekk_list(A=A,B=B,C=C,idmod=22,y.dim=2)
et_simulate.mgarch(model=model.bekk,n=n,n.start=200,rseed=1)

e1_et$et[,1]
e2_et$et[,2]
nstart_n-f-100

e_cbind(e1,e2)

tsplot(e1,xlab="time",ylab="spot returns")
tsplot(e2,xlab="time",ylab="futures returns")

e1f_e[-c(1:(n-f)),1]
e2f_e[-c(1:(n-f)),2]

e1_e[-c((n-f+1):n),1]
e2_e[-c((n-f+1):n),2]

stats_garch.stats(e[c(1:(n-f)),],max.lag=24)
stats

varini_var(e[(nstart:(n-f)),])
H_matrix(c(varini[1,1],varini[1,2],varini[2,1],varini[2,2]),nrow=2,ncol=2)

ABCCBA1_t(C)%*%C
ABCCBA2_t(A)%*%e[(n-f),]%*%t(e[(n-f),])%*%A
ABCCBA3_t(B)%*%H%*%B

H1_ABCCBA1+ABCCBA2+ABCCBA3
H11[i]_H1[1,1]
H12[i]_H1[1,2]
H22[i]_H1[2,2]

sumABCCBA1_matrix(c(0,0,0,0),2,2)

for (i in 2:f)
{
  ABCCBA1f_t(A)%*%ABCCBA1%*%A+t(B)%*%ABCCBA1%*%B
  ABCCBA2f_t(A)%*%ABCCBA2%*%A+t(B)%*%ABCCBA2%*%B
  ABCCBA3f_t(A)%*%ABCCBA3%*%A+t(B)%*%ABCCBA3%*%B
  sumABCCBA1f=ABCCBA1f
  ABCCBA2_ABCCBA2f
  ABCCBA3_ABCCBA3f
  Hfore_t(C)%*%C+sumABCCBA1f+ABCCBA2f+ABCCBA3f
  H11[i]_Hfore[1,1]
  H12[i]_Hfore[1,2]
  H22[i]_Hfore[2,2]
}

hfo_H12[f]/H22[f]
print("Forecasted hedge ratio is equal to")
hfo
var_var(e[(1:(n-f)),])
hmv_var[1,2]/var[2,2]
print("Minimum-variance hedge ratio is equal to")
hmv

for (i in (n-f+1):n)
{
H_t(C)*C+t(A)*t(e[(i-1),])%*t(e[(i-1),])%*A+t(B)%*H%*B
H11[i]_H[1,1]
H12[i]_H[1,2]
H22[i]_H[2,2]
}

H11_H11[-c(1:(n-f))]
H12_H12[-c(1:(n-f))]
H22_H22[-c(1:(n-f))]

hrt_H12/H22
tsplot(hrt,xlab="time",ylab="GARCH hedge ratio")

mean(hrt)
max(hrt)
min(hrt)
var(hrt)

for (i in 1:f)
{
cvtv[i]_H11[i]-2*hrt[i]*H12[i]+(hrt[i]^2)*H22[i]
cvmv[i]_H11[i]-2*hmv*H12[i]+(hmv^2)*H22[i]
cvna[i]_H11[i]-2*H12[i]+H22[i]
cvfo[i]_H11[i]-2*hfo*H12[i]+(hfo^2)*H22[i]
}

tsplot(cvtv-cvmv,xlab="time",ylab="tv-mv")
tsplot(cvtv-cvna,xlab="time",ylab="tv-na")
tsplot(cvtv-cvfo,xlab="time",ylab="tv-fo")
tsplot(cvmv-cvna,xlab="time",ylab="mv-na")
tsplot(cvmv-cvfo,xlab="time",ylab="mv-fo")
tsplot(cvna-cvfo,xlab="time",ylab="na-fo")

hcons[1:39]_hmv
hcons[40:58]_hfo
hcons[59:62]_hmv
hcons[63:100]_hfo

print("Hedging effectiveness for the naive hedge is equal to")
\[ 1 - \frac{\text{var}(e1f - e2f)}{\text{var}(e1f)} \]

print("Hedging effectiveness for the minimum-variance hedge is equal to")
1 - \frac{\text{var}(e1f - hmv*e2f)}{\text{var}(e1f)}

print("Hedging effectiveness for the GARCH hedge is equal to")
1 - \frac{\text{var}(e1f - hrt*e2f)}{\text{var}(e1f)}

print("Hedging effectiveness for the forecasted hedge is equal to")
1 - \frac{\text{var}(e1f - hfo*e2f)}{\text{var}(e1f)}

print("Hedging effectiveness for the alternative hedge is equal to")
1 - \frac{\text{var}(e1f - hcons*e2f)}{\text{var}(e1f)}
S-Plus program implemented in Example 2:

```r
e1_vector()
e2_vector()
e1f_vector()
e2f_vector()
H11_vector()
H12_vector()
H22_vector()
hrt_vector()
cvtv_vector()
cvmv_vector()
cvna_vector()
cvfo_vector()
cvco_vector()

var_matrix()
varini_matrix()

import.data(FileName="C:/My Documents/ThesisFinal/Chapters456/HeatingCrude.xls",DataFrame="HeatingCrude",FileType="Excel")

import.data(FileName="C:/My Documents/ThesisFinal/Chapters456/HChapter4.xls",DataFrame="H11H12H22",FileType="Excel")

tsplot(HeatingCrude[,4],xlab="time",ylab="heating oil prices (logarithms)"
tsplot(HeatingCrude[,5],xlab="time",ylab="crude oil prices (logarithms)"

tsplot(HeatingCrude[,6],xlab="time",ylab="heating oil returns"
tsplot(HeatingCrude[,7],xlab="time",ylab="crude oil returns"

n_nrow(HeatingCrude)
f_61

nstart_n-f-100

e_cbind(HeatingCrude[,6],HeatingCrude[,7])

e1f_e[-c(1:(n-f)),1]
e2f_e[-c(1:(n-f)),2]

e1_e[-c((n-f+1):n),1]
e2_e[-c((n-f+1):n),2]

stats_garch.stats(e[c(1:(n-f))],max.lag=24)
stats
```

250
varini_var(e[(nstart:(n-f))],)
H_matrix(c(varini[1,1],varini[1,2],varini[2,1],varini[2,2]),nrow=2,ncol=2)

hfo_0.75627
print("Forecasted hedge ratio is equal to")
hfo

var_var(e[(1:(n-f))],)
hmv_var[1,2]/var[2,2]
print("Minimum-variance hedge ratio is equal to")
hmv

H11_H11H12H22[,1]
H12_H11H12H22[,2]
H22_H11H12H22[,3]

hrt_H12/H22

tsplot(hrt,xlab="time",ylab="GARCH hedge ratio")

mean(hrt)
max(hrt)
min(hrt)
var(hrt)

for (i in 1:f)
{
    cvtv[i]_H11[i]-2*hrt[i]*H12[i]+(hrt[i]^2)*H22[i]
cvmv[i]_H11[i]-2*hmv*H12[i]+(hmv^2)*H22[i]
cvna[i]_H11[i]-2*H12[i]+H22[i]
cvfo[i]_H11[i]-2*hfo*H12[i]+(hfo^2)*H22[i]
}

var(cvtv-cvmv)
var(cvtv-cvna)
var(cvtv-cvfo)
var(cvmv-cvna)
var(cvmv-cvfo)
var(cvna-cvfo)

hcons[1:16]_hmv
hcons[17:40]_1
hcons[41:61]_hmv

print("Hedging effectiveness for the naive hedge is equal to")
1-var(e1f-e2f)/var(e1f)

print("Hedging effectiveness for the minimum-variance hedge is equal to")
251
1 - \text{var}(e_{1f} - \text{hmv}*e_{2f}) / \text{var}(e_{1f})

\text{print}("Hedging effectiveness for the forecasted hedge is equal to")
1 - \text{var}(e_{1f} - \text{hfo}*e_{2f}) / \text{var}(e_{1f})

\text{print}("Hedging effectiveness for the GARCH hedge is equal to")
1 - \text{var}(e_{1f} - \text{hrt}*e_{2f}) / \text{var}(e_{1f})

\text{print}("Hedging effectiveness for the alternative hedge is equal to")
1 - \text{var}(e_{1f} - \text{hcons}*e_{2f}) / \text{var}(e_{1f})
S-Plus program implemented in Example 3:

```s-plus
x_vector()
y_vector()
e_vector()
phi_vector()
theta_vector()
residuals_data.frame()
residualsmodified_data.frame()
Sigma_matrix()
A_matrix()
Astar1_matrix()
Astar2_matrix()
phimatrix_matrix()
thermatmatrix_matrix()
phimatrixinv_matrix()
C_matrix()
alpha_matrix()
beta_matrix()
psi_vector()
cv_vector()

set.seed(1)

x[1]_0
x[2]_0
x[3]_0
y[1]_0
y[2]_0
y[3]_0

e_rnorm(1200,0,1)

for (i in 4:1200)
{
x[i]_1.4*x[i-1]-0.2*x[i-2]-0.2*x[i-3]+e[i]
y[i]_1.2*y[i-1]-0.2*y[i-2]+e[i]
}

x_x[-c(1:200)]
y_y[-c(1:200)]

mle.x_arima.mle(x,model=list(order=c(2,1,0)))
diag.x_arima.diag(mle.x,resid=T,plot=F)
phi[1]_1-sum(mle.x$model$ar)
theta[1]_1-sum(mle.x$model$ma)
```

253
mle.y_arima.mle(y,model=list(order=c(1,1,0)))
diag.y_arima.diag(mle.y,resid=T,plot=F)
phi[2]_1-sum(mle.y$model$ar)
theta[2]_1-sum(mle.y$model$ma)
residuals_cbind(diag.x$resid,diag.y$resid)
residualsmodified_na.omit(residuals)

Sigma_var(residualsmodified)
print("Sigma")
Sigma

ev_eigen(Sigma)
print("ev")
ev
cvproportion_ev$values[2]/(ev$values[1]+ev$values[2])
print("cvproportion")
cvproportion

# A is a nonsingular matrix containing all the eigenvectors of Sigma
A_eigen(Sigma)$vectors
print("A")
A

# p is the number of variables in the system p_2
# r is the number of cointegrating vectors r_1
# Let A*=A^{-1}, then partition A* into two matrices A*1 and A*2
Astar1_A[,c(1:(p-r))]
Astar2_A[,c((p-r+1):p)]

phimatrix_diag(phi,nrow=p)
thetamatrix_diag(theta,nrow=p)
phimatrinxinv_solve(phimatrix)

C_phimatrixinv%*%thetamatrix%*%Astar1

alpha_-(C[[1:r],])
beta_t(C[[r+1]:p],)

psi_solve(beta,alpha)
print("psi")
psi

cv_x+psi*y
tsplot(cv,xlab="time",ylab="linear combination of cointegrating vector")
S-Plus program implemented in Example 4:

```r
x_vector()
y_vector()
e_vector()
phi_vector()
theta_vector()
residuals_data.frame()
residualsmodified_data.frame()
Sigma_matrix()
A_matrix()
Astar1_matrix()
Astar2_matrix()
phimatrix_matrix()
thetamatrix_matrix()
phimatrixinv_matrix()
C_matrix()
alpha_matrix()
beta_matrix()
psi_vector()
cv_vector()

set.seed(1)

x[1]_0
x[2]_0
x[3]_0
y[1]_0
y[2]_0
y[3]_0
e_rnorm(1200,0,1)

for (i in 4:1200)
{
x[i]_x[i-1]+e[i]-0.2*e[i-1]
y[i]_y[i-1]+sqrt(2)*e[i]+0.2*e[i-1]
}

x_x[-c(1:200)]
y_y[-c(1:200)]

mle.x_arima.mle(x,model=list(order=c(0,1,1)))
diag.x_arima.diag(mle.x,resid=T,plot=F)
phi[1]_1-sum(mle.x$model$ar)
theta[1]_1-sum(mle.x$model$ma)
```

255
mle_y_arima.mle(y, model=list(order=c(0,1,1)))
diag_y_arima.diag(mle_y, resid=T, plot=F)
phi[2]_1-sum(mle_y$model$ar)
theta[2]_1-sum(mle_y$model$ma)

residuals_cbind(diag_x$resid, diag_y$resid)
residualsmodified_na.omit(residuals)

Sigma_var(residualsmodified)
print("Sigma")
Sigma

ev_eigen(Sigma)
print("ev")
ev
cvproportion_ev$values[2]/(ev$values[1]+ev$values[2])
print("cvproportion")
cvproportion

# A is a nonsingular matrix containing all the eigenvectors of Sigma
A_eigen(Sigma)$vectors
print("A")
A

# p is the number of variables in the system p_2
# r is the number of cointegrating vectors r_1
# Let A**=Ainverse, then partition A* into two matrices A*1 and A*2
Astar1_A[,c(1:(p-r))]
Astar2_A[,c((p-r+1):p)]

phimatrix_diag(phi, nrow=p)
thetamatrix_diag(theta, nrow=p)
phimatrixinv_solve(phimatrix)

C_phimatrixinv%*%thetamatrix%*%Astar1
alpha_-(C[[1:r],])
beta_t(C[[r+1]:p],)

psi_solve(beta, alpha)
print("psi")
psi
cv_x+psi*y
tsplot(cv, xlab="time", ylab="linear combination of cointegrating vector")
S-Plus program implemented in Example 5:

```r
x_vector()
y_vector()
z_vector()
e1_vector()
e2_vector()
phi_vector()
theta_vector()
residuals_data.frame()
residualsmodified_data.frame()
Sigma_matrix()
A_matrix()
Astar1_matrix()
Astar2_matrix()
phimatrix_matrix()
thetamatrix_matrix()
phimatrixinv_matrix()
C_matrix()
alpha_matrix()
beta_matrix()
psi_vector()
cv_vector()

set.seed(1)

x[1]_0
x[2]_0
x[3]_0
y[1]_0
y[2]_0
y[3]_0
z[1]_0
z[2]_0
z[3]_0
e1_rnorm(1200,0,0.8)
e2_rnorm(1200,0,1)

for (i in 4:1200)
{
  x[i]_1.4*x[i-1]-0.2*x[i-2]-0.2*x[i-3]+e1[i]
y[i]_1.3*y[i-1]-0.3*y[i-2]+e1[i]+e2[i]
z[i]_1.6*z[i-1]-0.6*z[i-2]+e2[i]-0.8*e2[i-1]
}

x_x[-c(1:200)]
```

257
y_y[-c(1:200)]
z_z[-c(1:200)]

mle.x_arima.mle(x,model=list(order=c(2,1,0)))
diag.x_arima.diag(mle.x,resid=T,plot=F)
phi[1]_1-sum(mle.x$model$ar)
theta[1]_1-sum(mle.x$model$ma)

mle.y_arima.mle(y,model=list(order=c(1,1,0)))
diag.y_arima.diag(mle.y,resid=T,plot=F)
phi[2]_1-sum(mle.y$model$ar)
theta[2]_1-sum(mle.y$model$ma)

mle.z_arima.mle(z,model=list(order=c(1,1,1)))
diag.z_arima.diag(mle.z,resid=T,plot=F)
phi[3]_1-sum(mle.z$model$ar)
theta[3]_1-sum(mle.z$model$ma)

residuals_cbind(diag.x$resid,diag.y$resid,diag.z$resid)
residualsmodified_na.omit(residuals)

Sigma_var(residualsmodified)
print("Sigma")
Sigma

ev_eigen(Sigma)
print("ev")
ev
print("cvproportion")
cvproportion

# A is a nonsingular matrix containing all the eigenvectors of Sigma
A_eigen(Sigma)$vectors
print("A")
A

# p is the number of variables in the system p_3
# r is the number of cointegrating vectors r_1
# Let A*=Ainverse,then partition A* into two matrices A*1 and A*2
Astar1_A[,c(1:(p-r))]
Astar2_A[,c((p-r+1):p)]

phimatrix_diag(phi,nrow=p)
thetamatrix_diag(theta,nrow=p)
phimatrixinv_solve(phimatrix)
C_phimatrixinv%*%thetamatrix%*%Astar1

alpha_-(C[(1:r),])
beta_t(C[(r+1):p,])

psi_solve(beta,alpha)
print("psi")
psi

meanx_vector()
meanx[1]_mean(diff(x))
meanx[2]_mean(diff(y))
meanx[3]_mean(diff(z))

trend_vector()

for (t in 1:997)
{
}

cv_x+psi[1]*y+psi[2]*z
tspplot(cv,xlab="time",ylab="linear combination of cointegrating vector")
S-Plus program implemented in Example 6:

```splus
x_vector()
y_vector()
e_vector()
phi_vector()
theta_vector()
residuals_data.frame()
residualsmodified_data.frame()
Sigma_matrix()
A_matrix()
Astar1_matrix()
Astar2_matrix()
phimatrix_matrix()
thetamatrix_matrix()
phimatrixinv_matrix()
C_matrix()
alphamatrix()
betamatrix()
psivector()
rbccv_vector()
johcv_vector()

import.data(FileName="C:/My Documents/ThesisFinal/Chapters456/$USUK.xls",DataFrame="USUK",FileType="Excel")

x_log(USUK[,1])
y_log(USUK[,2])

tsplot(x,xlab="time",ylab="futures prices (logarithms)"

tsplot(y,xlab="time",ylab="spot prices (logarithms)"

mle.x_arima.mle(x,model=list(order=c(3,1,0)))
diag.x_arima.diag(mle.x,resid=T,plot=F)
phi[1]_1-sum(mle.x$model$ar)
theta[1]_1-sum(mle.x$model$ma)

mle.y_arima.mle(y,model=list(order=c(0,1,3)))
diag.y_arima.diag(mle.y,resid=T,plot=F)
phi[2]_1-sum(mle.y$model$ar)
theta[2]_1-sum(mle.y$model$ma)

residuals_cbind(diag.x resid,diag.y resid)
residualsmodified_na.omit(residuals)

Sigma_var(residualsmodified)
print("Sigma")
```

260
Sigma
ev_eigen(Sigma)
print("ev")
ev
cvproportion_ev$values[2]/(ev$values[1]+ev$values[2])
print("cvproportion")
cvproportion

# A is a nonsingular matrix containing all the eigenvectors of Sigma
A_eigen(Sigma)$vectors
print("A")
A

# p is the number of variables in the system p_2
# r is the number of cointegrating vectors r_1
# Let A*=Ainverse, then partition A* into two matrices A*1 and A*2
Astar1_A[,c(1:(p-r))]
Astar2_A[,c((p-r+1):p)]

phimatrix_diag(phi,nrow=p)
thetamatrix_diag(theta,nrow=p)
phimatrixinv_solve(phimatrix)

C_phimatrixinv%*%thetamatrix%*%Astar1

alpha_-(C[[1:r),]])
beta_t(C[[r+1):p,]])

psi_solve(beta,alpha)
print("psi")
psi

johcv_x-1.0118*y
rbccv_x-1.016763*y

tsplot(johcv,xlab="time",ylab="linear combination of cointegrating vector")
tsplot(rbccv,xlab="time",ylab="linear combination of cointegrating vector")
S-Plus program implemented in Example 7:

```r
x_vector()
y_vector()
z_vector()
phi_vector()
theta_vector()
residuals_data.frame()
residualsmodified_data.frame()
Sigma_matrix()
A_matrix()
Astar1_matrix()
Astar2_matrix()
phimatrix_matrix()
thetamatrix_matrix()
phimatrixinv_matrix()
C_matrix()
alpha_matrix()
beta_matrix()
psi_vector()
rbcv_vector()
johcv_vector()

import.data(FileName="C:/My Documents/ThesisFinal/Chapters456/$Asean.xls",DataFrame="Asean",FileType="Excel")

x_log(Asean[,2])
y_log(Asean[,3])
z_log(Asean[,4])

x_x[-c(2611:3174)]
y_y[-c(2611:3174)]
z_z[-c(2611:3174)]

tsplot(x,xlab="time",ylab="Malaysian Ringatt vs. US Dollar (logarithms)")
tsplot(y,xlab="time",ylab="Philippine Peso vs. US Dollar (logarithms)")
tsplot(z,xlab="time",ylab="Thai Baht vs. US Dollar (logarithms)")

mle.x_arima.mle(x,model=list(order=c(4,1,0)))
diag.x_arima.diag(mle.x,resid=T,plot=F)
phi[1]_1-sum(mle.x$model$ar)
theta[1]_1-sum(mle.x$model$ma)

mle.y_arima.mle(y,model=list(order=c(1,1,2)))
diag.y_arima.diag(mle.y,resid=T,plot=F)
phi[2]_1-sum(mle.y$model$ar)
theta[2]_1-sum(mle.y$model$ma)
```
mle.z_arima.mle(z, model=list(order=c(0,1,4)))
diag.z_arima.diag(mle.z, resid=T, plot=F)
phi[3]_1-sum(mle.z$model$ar)
theta[3]_1-sum(mle.z$model$ma)

residuals_{cbind(diag.x$resid, diag.y$resid, diag.z$resid)}
residuals_{modified_na.omit(residuals)}

Sigma_var(residuals_{modified})
print("Sigma")
Sigma

ev_eigen(Sigma)
print("ev")
ev

print("cvproportion")
cvproportion

# A is a nonsingular matrix containing all the eigenvectors of Sigma
A_eigen(Sigma)$vectors
print("A")
A

# p is the number of variables in the system p_3
# r is the number of cointegrating vectors r_1
# Let A*=A^{-1}, then partition A* into two matrices A*1 and A*2

Astar1_A[,c(1:(p-r))]
Astar2_A[,c((p-r+1):p)]

phimatrix_diag(phi, nrow=p)
thetamatrix_diag(theta, nrow=p)
phimatrixinv_solve(phimatrix)

C_phimatrixinv%*%thetamatrix%*%Astar1

alpha_-(C[[1:r],])
beta_t(C[[r+1]:p,])

psi_solve(beta, alpha)
print("psi")
psi

meanx_vector()
meanx[1].mean(diff(x))
meanx[2]_mean(diff(y))
meanx[3]_mean(diff(z))

trend_vector()

for (t in 1:2610)
{
}

rbccv_x+psi[1]*y+psi[2]*z-trend
johcv_x+28.229*y+298.19*z+trend

tsplot(rbccv,xlab="time",ylab="linear combination of cointegrating vector")
tsplot(johcv,xlab="time",ylab="linear combination of cointegrating vector")
S-Plus program implemented in Example 8:

delta_vector()
H_vector()
e_vector()
sp_vector()
z_vector()
fu_vector()
e1_vector()
e2_vector()
e1f_vector()
e2f_vector()
H11_vector()
H12_vector()
H22_vector()
hrt_vector()
cvtv_vector()
cvmv_vector()
cvna_vector()
e1e2_data.frame()
var_matrix()

set.seed(1)

n_1210
f_200
sd_sqrt(0.5)
beta_1
alpha_0.6

c_0
a_0.2
b_0.8
d_0.1

delta_rnorm(n,sd)
H[1]_0
e[1]_0
sp[1]_1000
z[1]_0
fu[1]_(1/beta)*(sp[1]-z[1])

n1_25

for (i in 2:n)
{

\[ H[i] = c + a \cdot e[i-1]^2 + b \cdot H[i-1] + d \cdot z[i-1]^2 \]
\[ t_{rt}(1,n1) \]
\[ e[i] = \sqrt{H[i] \cdot (n1-2)/n1} \cdot t \]
\[ sp[i] = sp[i-1] + e[i] \]
\[ z[i] = \delta[i] + \alpha \cdot \delta[i-1] \]
\[ fu[i] = (1/\beta) \cdot (sp[i] - z[i]) \]
\}

\[ sp_{sp}[-c(1:10)] \]
\[ fu_{fu}[-c(1:10)] \]
\[ e1\_diff(sp) \]
\[ e2\_diff(fu) \]
\[ tsplot(sp, xlab="time", ylab="spot prices") \]
\[ tsplot(fu, xlab="time", ylab="futures prices") \]
\[ tsplot(e1, xlab="time", ylab="spot returns") \]
\[ tsplot(e2, xlab="time", ylab="futures returns") \]

\[ \text{n\_length}(sp) \]
\[ e1\_insert.row(e1, 1, 1) \]
\[ e2\_insert.row(e2, 1, 1) \]
\[ e1[1] = c(0) \]
\[ e2[1] = c(0) \]
\[ e1e2\_cbind(e1, e2) \]
\[ e1f\_e1e2[-c(1:(n-f)), 1] \]
\[ e2f\_e1e2[-c(1:(n-f)), 2] \]
\[ e1\_ele2[-c((n-f+1):n), 1] \]
\[ e2\_ele2[-c((n-f+1):n), 2] \]
\[ \text{stats\_garch.stats}(e1e2[c(1:(n-f)), ], max.lag=24) \]
\[ \text{stats} \]
\[ \text{var\_var}(e1e2[(1:(n-f)),]) \]
\[ \text{hmv\_var}[1,2]/\text{var}[2,2] \]
\[ \text{print("Minimum-variance hedge ratio is equal to")} \]
\[ \text{hmv} \]
\[ \text{for} (i \text{ in} (n-f+1):n) \]
\[ \{ \]
\[ H11[i] = c + a \cdot e[i-1]^2 + b \cdot H[i-1] + d \cdot z[i-1]^2 \]
\[ H12[i] = H11[i]/\beta \]
\[ H22[i] = ((z[i-1] - \alpha \cdot \delta[i-1])^2 + H11[i] + \text{sd}^2)/\beta^2 \]
\}
H11_H11[-c(1:(n-f))]  
H12_H12[-c(1:(n-f))]  
H22_H22[-c(1:(n-f))]  

hrt_H12/H22  

tsplot(hrt,xlab="time",ylab="GARCH-X hedge ratio")  

mean(hrt)  
max(hrt)  
min(hrt)  
var(hrt)  

for (i in 1:f)  
{  
cvtv[i]_H11[i]-2*hrt[i]*H12[i]+(hrt[i]^2)*H22[i]  
cvmv[i]_H11[i]-2*hmv*H12[i]+(hmv^2)*H22[i]  
cvna[i]_H11[i]-2*H12[i]+H22[i]  
}  

tsplot(cvtv-cvmv,xlab="time",ylab="tv-mv")  
tsplot(cvtv-cvna,ylab="tv-na")  
tsplot(cvmv-cvna,xlab="time",ylab="mv-na")  

print("Hedging effectiveness for the naive hedge is equal to")  
1-var(e1f-e2f)/var(e1f)  

print("Hedging effectiveness for the minimum-variance hedge is equal to")  
1-var(e1f-hmv*e2f)/var(e1f)  

print("Hedging effectiveness for the GARCH hedge is equal to")  
1-var(e1f-hrt*e2f)/var(e1f)
S-Plus program implemented in Example 9:

```r
# Import data from HeatingCrude.xls
import.data(FileName="C:/My Documents/ThesisFinal/Chapters456/$HeatingCrude.xls",DataFrame="HeatingCrude",FileType="Excel")

# Import data from H11H12H22.xls
import.data(FileName="C:/My Documents/ThesisFinal/Chapters456/$HChapter6.xls",DataFrame="H11H12H22",FileType="Excel")

tsp.plot(HeatingCrude[,4],xlab="time",ylab="heating oil prices (logarithms)")
tsp.plot(HeatingCrude[,5],xlab="time",ylab="crude oil prices (logarithms)")
tsp.plot(HeatingCrude[,6],xlab="time",ylab="heating oil returns")
tsp.plot(HeatingCrude[,7],xlab="time",ylab="crude oil returns")

n_nrow(HeatingCrude)
f_61

nstart_n-f-100

e_cbind(HeatingCrude[,6],HeatingCrude[,7])

e1f_e[-c(1:(n-f)),1]
e2f_e[-c(1:(n-f)),2]

e1_e[-c((n-f+1):n),1]
e2_e[-c((n-f+1):n),2]

stats_garch.stats(e[c(1:(n-f)),],max.lag=24)
stats

varini_var(e[(nstart:(n-f))])
```

268
H_matrix(c(varini[1,1],varini[1,2],varini[2,1],varini[2,2]),nrow=2,ncol=2)

var_var(e[(1:(n-f)),])

hmv_var[1,2]/var[2,2]

print("Minimum-variance hedge ratio is equal to")

hmv

H11_H11H12H22[,1]
H12_H11H12H22[,2]
H22_H11H12H22[,3]

hrt_H12/H22

tsplot(hrt,xlab="time",ylab="GARCH-X hedge ratio")

mean(hrt)

max(hrt)

min(hrt)

var(hrt)

for (i in 1:f)
{
  cvtv[i]_H11[i]-2*hrt[i]*H12[i]+(hrt[i]^2)*H22[i]
  cvmv[i]_H11[i]-2*hmvr12[i]+(hmvr2)*H22[i]
  cvna[i]_H11[i]-2*H12[i]+H22[i]
}

tsplot(cvtv-cvmv,xlab="time",ylab="tv-mv")

tsplot(cvtv-cvna,xlab="time",ylab="tv-na")

tsplot(cvmv-cvna,xlab="time",ylab="mv-na")

hcons[1:16]_hmv
hcons[17:40]_1
hcons[41:61]_hmv

print("Hedging effectiveness for the naive hedge is equal to")

1-var(e1f-e2f)/var(e1f)

print("Hedging effectiveness for the minimum-variance hedge is equal to")

1-var(e1f-hmv*e2f)/var(e1f)

print("Hedging effectiveness for the GARCH-X hedge is equal to")

1-var(e1f-hrt*e2f)/var(e1f)
RATS program implemented in Example 2:

Note the form of the conditional variance/covariance matrix $H$ in the RATS code is specified as $H(t)=C'C+B'e(t-1)e(t-1)'B+A'H(t-1)A$. In S-Plus and in the thesis $H(t)$ was specified to be equal to $H(t)=C'C+A'e(t-1)e(t-1)'A+B'H(t-1)B$.

* GARCHMV.PRG
* Updated, March 2003 to include dynamic conditional correlation
* Also, all models have been rewritten to stuff the $uu'$ into a
* matrix of series. This allows a bit more flexibility in handling
* both initial conditions, and parameterizing the functions.
*
* CALENDER(IRREGULAR) ALLOCATE 1313 OPEN DATA HEATINGCRUDE.RAT
DATA(FORMAT=RATS) / DATE HEAT CRUDE LNHEAT LNCRUDE DLNHEAT DLNCRUDE
*
* Starting and ending points for GARCH estimation. GSTART must
* be at least entry 2.
*
COMPUTE GSTART=3, GEND=1252
*
* Parameters for the regression function
*
DECLARE VECTOR[SERIES] Y(1) U(1) DECLARE VECTOR[FRML] RESID(1)
DECLARE VECTOR[SERIES] Y(2) U(2) DECLARE VECTOR[FRML] RESID(2)

SET Y(1) = DLNHEAT
SET Y(2) = DLNCRUDE

NONLIN(PARMSET=MEANPARMS) B11 B21
FRML RESID(1) = (Y(1)-B11)
FRML RESID(2) = (Y(2)-B21)

* Do initial regression. Copy initial values for regression parameters
*
LINREG Y(1) / U(1) # CONSTANT
COMPUTE B11 = %BETA(1)
LINREG Y(2) / U(2) # CONSTANT
COMPUTE B21 = %BETA(1)

* * Get the covariance matrix of the residuals.
* *
VCV(MATRIX=RR,NOPRINT) # U(1) U(2)

* * h will have the sequence of variance estimates
* uu will have the sequence of uu' matrices
* *
DECLARE SYMMETRIC[SERIES] H(2,2)
DECLARE SYMMETRIC[SERIES] UU(2,2)

* * hx and uux are used when extracting elements from h and uu.
* ux is used when extracting a u vector
* *
DECLARE SYMMETRIC HX(2,2) UUX(2,2)
DECLARE VECTOR UX(2)

* * This is used to initialize pre-sample variances.
* If you want the pre-sample uu' to be the unconditional variance,
* change the right side of the set uu(i,j) to rr(i,j) (same as h).
* *
DO I=1,2
    DO J=1,I
    SET H(I,J) = RR(I,J)
    SET UU(I,J) = RR(I,J)
    END DO J
END DO I

* * This is a standard log likelihood formula for any bivariate
* ARCH, GARCH, ARCH-M,... The difference among these will be in
* the definitions of HF and RESID. The function %XT pulls information
* out of a matrix of SERIES, while %PT puts information into one.
* *
DECLARE FRML[SYMMETRIC] HF FRML LOGL = $
    HX = HF(T) , $ 
    UX = %XT(U,T) , UUX = %OUTERXX(UX),$
    %PT(H,T,HX),%PT(UU,T,%OUTERXX(UX)),$

271
%LOGDENSITY(HX,UX)

* Positive definite parameterization (BEKK,EK)
* This enforces a positive definite covariance matrix by writing the
  covariance matrix evolution as
* \[ V(t) = C'C + B'u(t)u(t)'B + A'V(t-1)A \]
* Note that the parameters are not globally identified: changing the
  signs of all members of C,B or A will have no effect on the function
  value. Using METHOD=SIMPLEX to begin is quite important with this
  setup, to pull the estimates away from zero before starting the
  derivative-based methods.
*
DECLARE RECTANGULAR VAR(2,2)
COMPUTE VAR = ||VAR(1,1), VAR(1,2) | VAR(2,1), VAR(2,2)||
DECLARE RECTANGULAR VBR(2,2)
COMPUTE VBR = ||VBR(1,1), VBR(1,2) | VBR(2,1), VBR(2,2)||
DECLARE RECTANGULAR VCR(2,2)
COMPUTE VCR = ||VCR(1,1), VCR(1,2) | 0, VCR(2,2)||

NONLINEAR(PARMSET=GARCHPARMS) VAR VBR VCR VAR(1,2)=VAR(2,1)=VBR(1,2)=VBR(2,1)=VCR(2,1)=0

FRML HF = $\left( HX=%XT(H,T-1)),(UUX=%XT(UU,T-1)),$
  \%INNERXX(VCR)+\%MQFORM(HX,VAR)+\%MQFORM(UUX,VBR) \right.$

* Initialize c’s from the decomp of the covariance matrix
*
COMPUTE VCR = \%DECOMP(RR)
COMPUTE VAR(1,1)=VAR(2,2)=0.95, VBR(1,1)=VBR(2,2)=0.25
COMPUTE VAR(1,2)=VAR(2,1)=VBR(1,2)=VBR(2,1)=0

MAXIMIZE(PARMSET=MEANPARMS+GARCHPARMS,METHOD=SIMPLEX,ITERS=30)
  LOGL GSTART GEND
MAXIMIZE(PARMSET=MEANPARMS+GARCHPARMS,METHOD=BFGS,ITERS=100)
  LOGL GSTART GEND

* Forecasting
*
DO TIME=GEND+1,GEND+61
Compute $H_x = HF(TIME), \%PT(UU, TIME, HX), \%PT(H, TIME, HX)$

End do TIME

Compute $H_{FO} = H(2,1)(1313)/H(2,2)(1313)$
Display 'FORECASTED HEDGE RATIO IS EQUAL TO' $H_{FO}$

VCV(MATRIX=RR, PRINT) GSTART GEND # DLNHEAT DLNCRUDE
Compute $MVHR = RR(1,2)/RR(2,2)$
Display 'MINIMUM-VARIANCE HEDGE RATIO IS EQUAL TO' $MVHR$

Statistics(NOPRINT) DLNHEAT GEND+1 GEND+61
Compute $VARUNH = \%VARIANCE$

Smpl 1253 1313

Set $NA = DLNHEAT-DLNCRUDE$ Statistics(NOPRINT) NA
Compute $VARNAH = \%VARIANCE$
Display 'VARIANCE OF THE NAIVE HEDGE PORTFOLIO' $VARNAH$

Set $MV = DLNHEAT-MVHR*DLNCRUDE$ Statistics(NOPRINT) MV
Compute $VARMVH = \%VARIANCE$
Display 'VARIANCE OF THE MINIMUM-VARIANCE HEDGE PORTFOLIO' $VARMVH$

Set $FO = DLNHEAT-H_{FO}*DLNCRUDE$ Statistics(NOPRINT) FO
Compute $VARFOH = \%VARIANCE$
Display 'VARIANCE OF THE FORECASTED HEDGE PORTFOLIO' $VARFOH$

Display 'HEDGING EFFECTIVENESS FOR NAIVE HEDGE IS EQUAL TO' $1-VARNAH/VARUNH$
Display 'HEDGING EFFECTIVENESS FOR MINIMUM-VARIANCE HEDGE IS EQUAL TO' $1-VARMVH/VARUNH$
Display 'HEDGING EFFECTIVENESS FOR FORECASTED HEDGE IS EQUAL TO' $1-VARFOH/VARUNH$

Declare rectangular $HT(2,2)$
Compute $HT = RR$

Do Step=0,60
Compute TIME = GEND+STEP
Maximize(PARMSET=MEANPARMS+GARCHPARMS, METHOD=BFGS, ITERS=100)$
Logl GSTART TIME
Compute $HT = TR(VCR) * VCR + TR(VAR) * HT * VAR + TR(VBR) * (%)XT(U, TIME)) * TR(\%XT(U, TIME))*VBR$
SetHT11 TIME+1 TIME+1 = HT(1,1)
Set HT12 TIME+1 TIME+1 = HT(1,2)
Set HT22 TIME+1 TIME+1 = HT(2,2)
SetHR TIME+1 TIME+1 = HT(1,2)/HT(2,2)
End do Step Statistics HR

273
SET GA = DLNHEAT-HR*DLNCRUDE STATISTICS(NOPRINT) GA
COMPUTE VARGA = %VARIANCE
DISPLAY 'VARIANCE OF THE TIME-VARYING (GARCH) HEDGE PORTFOLIO'
VARGA DISPLAY 'TIME-VARYING (GARCH) HEDGING EFFECTIVENESS IS EQUAL TO'$
1-VARGA/VARUNH
RATS program implemented in Example 9:

Note the form of the conditional variance/covariance matrix $H$ in the RATS code is specified as $H(t)=C'C+B'e(t-1)e(t-1)'B+A'H(t-1)A$. In S-Plus and in the thesis $H(t)$ was specified to be equal to $H(t)=C'C+A'e(t-1)e(t-1)'A+B'H(t-1)B$.

* * GARCHMV.PRG
* Updated, March 2003 to include dynamic conditional correlation
* Also, all models have been rewritten to stuff the $uu'$ into a
* matrix of series. This allows a bit more flexibility in handling
* both initial conditions, and parameterizing the functions.
*
CALENDER(IRREGULAR)
ALLOCATE 1313
OPEN DATA HEATINGCRUDE.RAT
DATA(FORMAT=RATS) / DATE HEAT CRUDE LNHEAT LNCRUDE DLNHEAT
DLNCRUDE SET CV = LNHEAT-0.88466*LNCRUDE

* * Starting and ending points for GARCH estimation. GSTART must
* be at least entry 2.
*
COMPUTE GSTART=3, GEND=1252

* Parameters for the regression function
*
DECLARE VECTOR[SERIES] Y(1) U(1) DECLARE VECTOR[FRML] RESID(1)
DECLARE VECTOR[SERIES] Y(2) U(2) DECLARE VECTOR[FRML] RESID(2)

SET Y(1) = DLNHEAT
SET Y(2) = DLNCRUDE

NONLIN(PARMSET=MEANPARMS) B11 B21
FRML RESID(1) = (Y(1)-B11)
FRML RESID(2) = (Y(2)-B21)

* * Do initial regression. Copy initial values for regression parameters
*
LINREG Y(1) / U(1) # CONSTANT
COMPUTE B11 = %BETA(1)
LINREG Y(2) / U(2) # CONSTANT
COMPUTE B21 = %BETA(1)

* Get the covariance matrix of the residuals.
*
VCV(MATRIX=RR,NOPRINT) # U(1) U(2)

* h will have the sequence of variance estimates
* uu will have the sequence of uu' matrices
*
DECLARE SYMMETRIC[SERIES] H(2,2)
DECLARE SYMMETRIC[SERIES] UU(2,2)

* hx and uux are used when extracting elements from h and uu.
* ux is used when extracting a u vector
*
DECLARE SYMMETRIC HX(2,2) UUX(2,2)
DECLARE VECTOR UX(2)

* This is used to initialize pre-sample variances.
* If you want the pre-sample uu' to be the unconditional variance,
* change the right side of the set uu(i,j) to rr(i,j) (same as h).
*
DO I=1,2
   DO J=1,I
      SET H(I,J) = RR(I,J)
      SET UU(I,J) = RR(I,J)
   END DO J
END DO I

* * This is a standard log likelihood formula for any bivariate
* ARCH, GARCH, ARCH-M,... The difference among these will be in
* the definitions of HF and RESID. The function %XT pulls information
* out of a matrix of SERIES, while %PT puts information into one.
*
DECLARE FRML[SYMMETRIC] HF
FRML LOGL = $ 
HX = HF(T), $ 
UX = %XT(U,T), UUX = %OUTERXX(UX),$ 
%PT(H,T,HX),%PT(UU,T,%OUTERXX(UX)),$ 
%LOGDENSITY(HX,UX) 
* 
Positive definite parameterization (BEKK,EK) 
* This enforces a positive definite covariance matrix by writing the 
covariance matrix evolution as 
* 
H(t) = C'C + B'u(t)u(t)'B + A'H(t-1)A 
* 
Note that the parameters are not globally identified: changing the 
* signs of all members of C,B or A will have no effect on the function 
* value. Using METHOD=SIMPLEX to begin is quite important with this 
* setup, to pull the estimates away from zero before starting the 
* derivative-based methods. 
* 
DECLARE RECTANGULAR VAR(2,2) 
COMPUTE VAR = ||VAR(1,1), VAR(1,2) | VAR(2,1), VAR(2,2)|| 
DECLARE RECTANGULAR VBR(2,2) 
COMPUTE VBR = ||VBR(1,1), VBR(1,2) | VBR(2,1), VBR(2,2)|| 
DECLARE RECTANGULAR VCR(2,2) 
COMPUTE VCR = ||VCR(1,1), VCR(1,2) | 0, VCR(2,2)|| 
DECLARE RECTANGULAR VDR(1,2) 
COMPUTE VDR = ||VDR(1,1), VDR(1,2)|| 
NONLINEAR(PARMSET=GARCHPARMS) VAR VBR VCR VDR VAR(1,2)=VAR(2,1)$ 
=VBR(1,2)=VBR(2,1)=VCR(2,1)=0 
FRML HF = $ 
(HX=%XT(H,T-1)),(UUX=%XT(UU,T-1)),$ 
%INNERXX(VCR)+%MQFORM(HX,VAR)+%MQFORM(UUX,VBR)+CV(T-1)**2*(TR(VDR)*VDR) 
* 
Initialize c’s from the decomp of the covariance matrix 
* 
COMPUTE VCR = %DECOMP(RR) 
COMPUTE VAR(1,1)=VAR(2,2)=0.95, VBR(1,1)=VBR(2,2)=0.25,$ 
VDR(1,1)=VDR(1,2)=0.001,VAR(1,2)=VAR(2,1)=VBR(1,2)=VBR(2,1)=0 
MAXIMIZE(PARMSET=MEANPARMS+GARCHPARMS,METHOD=SIMPLEX,ITERS=30)$ 
LOGL GSTART GEND 
MAXIMIZE(PARMSET=MEANPARMS+GARCHPARMS,METHOD=BFGS,ITERS=100)$ 
LOGL GSTART GEND
VCV(MATRIX=RR,PRINT) GSTART GEND # DLNHEAT DLNCRUDE
COMPUTE MVHR = RR(1,2)/RR(2,2)
DISPLAY ‘MINIMUM-VARIANCE HEDGE RATIO IS EQUAL TO’ MVHR

STATISTICS(NOPRINT) DLNHEAT GEND+1 GEND+61
COMPUTE VARUNH = %VARIANCE

SMPL 1253 1313

SET NA = DLNHEAT-DLNCRUDE
STATISTICS(NOPRINT) NA
COMPUTE VARNAH = %VARIANCE
DISPLAY ‘VARIANCE OF THE NAIVE HEDGE PORTFOLIO’ VARNAH

SET MV = DLNHEAT-MVHR*DLNCRUDE
STATISTICS(NOPRINT) MV
COMPUTE VARMVH = %VARIANCE
DISPLAY ‘VARIANCE OF THE MINIMUM-VARIANCE HEDGE PORTFOLIO’ VARMVH

DISPLAY ‘HEDGING EFFECTIVENESS FOR NAIVE HEDGE IS EQUAL TO’
1-VARNAH/VARUNH
DISPLAY ‘HEDGING EFFECTIVENESS FOR MINIMUM-VARIANCE HEDGE IS EQUAL TO’$
1-VARMVH/VARUNH

DECLARE RECTANGULAR HT(2,2)
COMPUTE HT = RR

DO STEP=0,60
  COMPUTE TIME = GEND+STEP
  MAXIMIZE(PARMSET=MEANPARMS+GARCHPARMS,METHOD=BFGS,ITERS=100)$
  LOGL GSTART TIME
  COMPUTE HT = TR(VCR)*VCR+TR(VAR)*HT*VAR+TR(VBR)*(%XT(U,TIME))*TR(%XT(U,TIME))*VBR+CV(TIME)**2*(TR(VDR)*VDR)
  SET HT11 TIME+1 TIME+1 = HT(1,1)
  SET HT12 TIME+1 TIME+1 = HT(1,2)
  SET HT22 TIME+1 TIME+1 = HT(2,2)
  SET HR TIME+1 TIME+1 = HT(1,2)/HT(2,2)
END DO

OPEN COPY HCHAPTER6.XLS
COPY(DATES,FORMAT=XLS,ORG=COLS) / HT11 HT12 HT22

SET GA = DLNHEAT-HR*DLNCRUDE STATISTICS(NOPRINT) GA
COMPUTE VARGA = %VARIANCE
DISPLAY ‘VARIANCE OF THE TIME-VARYING (GARCH) HEDGE PORTFOLIO’ VARGA

278
DISPLAY 'TIME-VARYING (GARCH-X) HEDGING EFFECTIVENESS IS EQUAL TO'
1 - VARGA/VARUNH