Modelling Intraday Stock Price Dynamics on the Malaysian Stock Exchange

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

By,

MOHD NIZAL HANIFF
CIMA (UK) Passed Finalist
Master of Science (Finance) (Universiti Putra Malaysia)

Accounting & Finance Section
Cardiff Business School
July 2006

ABSTRACT

The introduction of the Autoregressive Conditional Heteroskedasticity (ARCH) model in 1982 by Engle revolutionized the econometric treatment of volatility. The Generalized ARCH (GARCH) model and its variants have proved to be useful in capturing stylised facts about financial markets, which include volatility clustering, leptokurtosis in the distribution of returns, mean reversion tendencies and leverage effects. The Periodic GARCH (PGARCH) variants proposed by Bollerslev and Ghysels (1996), in particular, made it possible to explicitly incorporate the effects of periodicity in financial time series into the parameters of the volatility models. An investigation of return volatility using high frequency Kuala Lumpur Composite Index (KLCI) returns data shows that the intraday volatility pattern follows the double U-shaped pattern, which is consistent with the findings of other studies on markets that are closed during the lunch hour. The study also investigates the best technique for modelling and forecasting the intraday periodicity on the Kuala Lumpur Stock Exchange (KLSE), using both the jointly estimated and the two-step filtration approaches with different PGARCH structures. The results indicate that the PGARCH models produce superior model fit, better forecasting performances and superior forecast quality than the standard GARCH equivalents. However, the results suggest that Value-at-Risk (VaR) models, constructed from the PGARCH forecasts, produce poor results. This study also investigates the integrated realized volatility measure introduced by Andersen and Bollerslev (1998a), which can be constructed by summing up intraday squared returns. The results suggest that the daily integrated realized volatilities constructed using different intraday return sampling frequencies, produce superior forecasting performances for the GARCH models when compared with the results of the same models using the daily squared returns. The VaR models constructed from the GARCH forecasts and the Autoregressive and Moving Average (ARMA) forecasts appear to satisfy the requirements of the framework for interval forecast evaluation.

ACKNOWLEDGEMENTS

First and foremost, all praised be Allah (S.W.T) for giving me the excellent opportunity to pursue this Ph.D. degree, and without whose blessings I would not have been able to complete this thesis.

I would like to express my most heartfelt gratitude to my thesis supervisor, Dr Nick Taylor, the Distinguished Senior Research Fellow at the Cardiff Business School, for his effort, willingness, patience, invaluable ideas, direction on original research, consistent guidance, continual reading of my thesis at various stages and support in the preparation of this work.

I am grateful to the Universiti Teknologi MARA Malaysia (UiTM) and the Government of Malaysia for providing the financial support that allowed me to undertake this Ph.D., without which I would have had a difficult time continuing the lengthy process of Ph.D. research. I am also grateful to Cardiff Business School for the warmth and ever-helpful staff and the excellent facilities provided, making my stay here in Cardiff a very pleasant experience.

I would like to take this opportunity to register my most heartfelt gratitude to my father Haji Haniff Mohd Baki and my parents-in-law, Haji Mohd Drus Kasiman and Hajjah Hasnah Ali, as well as my two brothers, Hasnul Haniff and Anuar Haniff and the rest of my family members for their never-ending support, continual encouragement and prayers.

My very special thanks go to my beloved family here with me in Cardiff: to the most special person in my life, my beautiful and beloved wife, Suzana Mohd Drus for her sacrifices, understanding, care, support and patience and never-ending belief in my capability and talent. My special appreciation also goes to my children, Farah, Hariz and Sarah, who have brought me bundles of joy and made my life more bearable during the difficult phases in the preparation of this thesis. With their love and support, the writing and completion of this thesis have become a reality.

Last but not least, I would like to extend my appreciation to my fellow colleagues and friends here at the Cardiff Business School for their continual encouragement and selfless support in whatever problems I may have endured.

From the bottom of my heart, I would like to thank all of you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AIC	-	Akaike Information Criterion
AMEX	-	American Stock Exchange
AMV	-	Aggregate Market Value
ARCH	-	Autoregressive Conditional Heteroskedasticity
ARFIMA	-	Autoregressive Fractionally Integrated Moving Average
ARMA	-	Autoregressive and Moving Average
BFR	-	Broker Front End
CAC 40	-	Compagnie des Agents de Change 40 Index
CDS	-	Central Depository System
CI	-	Composite Index
DAX	-	Deutsche Aktienindex
DM	-	Deutsche Mark
EGARCH	-	Exponential GARCH
EWMA	-	Exponentially Weighted Moving Average
FFF	-	Flexible Fourier Form
FTA	-	Financial Times All Share
FTSE 100	-	Financial Times Stock Exchange 100 Index
GARCH	-	Generalized Autoregressive Conditional Heteroskedasticity
GARCH–M	-	GARCH in Mean
GJR-GARCH	-	Glosten, Jagannathan and Runkle GARCH
HISVOL	-	Historical Volatility
HLN	-	Harvey, Stephen and Newbold
ICT	-	Information and Communication Technology
ISD	-	Option-implied Volatility
KLCI	-	Kuala Lumpur Composite Index
KLSE	-	Kuala Lumpur Stock Exchange
KLSEB	-	Kuala Lumpur Stock Exchange Berhad
LIFFE	-	London International Financial Futures and Options Exchange
LL	-	Log Likelihood
MAFE	-	Mean Absolute Forecast Error
MCD	-	Malaysian Central Depository
MESDAQ	-	Malaysian Exchange of Securities Dealing and Automated Quotation
MLE	-	Maximum Likelihood Estimate
MSE	-	Malayan Stock Exchange
MSFE	-	Mean Squared Forecast Error
NASDAQ	-	National Association of Securities Dealers Automated Quotation
NSE	-	National Stock Exchange Mumbai
NYSE		New York Stock Exchange
OLS	-	Ordinary Least Squares
PGARCH	-	Periodic GARCH
QGARCH	-	Quadratic GARCH
QML	-	Quasi-Maximum Likelihood
RSS	-	Rolling Settlement System
S&P 100	-	Standard and Poor's 100 Index

SCANS	-	Securities Clearing Automated Network System
SIC	-	Schawrz Information Criterion
SCORE	-	System on Computerised Order Routing and Execution
SEM	-	Stock Exchange of Malaysia
SEMS	-	Stock Exchange of Malaysia and Singapore
SES	-	Stock Exchange of Singapore
SV	-	Stochastic Volatility
TGARCH	-	Threshold GARCH
UK	-	United Kingdom
US	-	United States
USD	-	US Dollar
VaR	-	Value-at-Risk

CHAPTER 1

INTRODUCTION

1.0 Research Background

Early econometric models and conventional times-series models operate under the assumption of constant variance. In 1982, Engle introduced the Autoregressive Conditional Heteroskedasticity (ARCH) model (an associated estimation methodology), which allows the variance to change over time as a function of past errors. This method has been found to be successful in modelling various times series applications because it is able to parameterise some of the stylised facts underlying financial markets. These stylised facts include volatility clustering (Mandelbrot, 1963), fat tails in the unconditional distribution, mean reversion and a phenomenon referred to as the leverage effect (Black, 1976). Since then, there has been a veritable explosion of papers analysing models of changing volatility. For example, a survey paper by Bollerslev, Chou and Kroner (1992) lists more than 100 papers on this subject. Some of the more popular variants of models of changing volatility have proved to be the variants of the Generalised ARCH (GARCH) model.

The GARCH class models have been proven valuable in modelling and forecasting the returns volatility at the monthly (see, for example, Cao and Tsay, 1992), weekly (see, for example, Cumby, Figlewski and Hasbrouck, 1993) and daily (see, for example, Akigray, 1989, Bera and Higgins, 1997) frequencies. Poon and Granger (2005) find that the GARCH models are as good as or even better than some of the traditional forecasting methods widely used previously. But with the availability of databases that provide tick data at the intraday frequency, serious doubt is cast upon the GARCH models regarding their ability to account adequately for intraday periodic effects that are often observed in financial time series. One of the most common periodic observations is a phenomenon which is referred to as the U-shaped pattern. Wood, McInish and Ord (1985) were among the earliest researchers to document the existence of a distinct U-shaped pattern in the variances of stock returns over the course of a trading day in the US markets. Similar patterns have also been observed in other markets in the US, such as the foreign exchange markets (see, for example, Baillie and Bollerslev, 1991, and Dacorogna, Muller, Nagler, Olsen and Pictet, 1993). This periodic pattern is now accepted as a typical feature of financial asset returns and recent studies have in fact shown that for markets that are closed during the lunch hour, a double U-shaped pattern is observed (see, for example, Taylor, 2004).

Although it is argued that this intraday periodicity is irrelevant for the analysis of data recorded on a daily, weekly or monthly basis, there are many recent studies that make use of intraday data in order to get a better appreciation of the interrelationship between different variables of interest during a particular trading day. For example, there are studies concerning the lead-lag relations between two or more markets that trade simultaneously.[1] Other examples include studies that explore the role of information flow and microstructure variables as determinants of intraday return volatility.[2] Therefore, it is not surprising to see why high frequency data has become very popular in many studies concerning volatility modelling and forecasting.

However, some recent studies focusing on intraday return volatility report that standard GARCH models, even though they represent the dominant technique for the

[1] See, for example, Baillie and Bollerslev (1991), who utilized hourly observations on five exchange rates and Chan, Chan and Karoyli (1991), who investigated five-minute returns, associated with stock index and stock index futures.

[2] See, for example, Bollerslev and Domowitz (1993), who analyzed five-minute foreign-exchange returns, and Locke and Sayers (1993), who modelled one-minute stock index futures returns.

empirical modelling of volatility at the daily (and lower) frequencies, are not able to capture adequately the systematic patterns observed during the trading day. Subsequently, many studies comparing the GARCH models with competing volatility models often report that the GARCH models tend to exhibit inferior forecasting performance (see Andersen and Bollerslev, 1997, and Martens, 2001). Andersen and Bollerslev (1997) argue that the pervasive intraday periodicity in the return volatility in foreign exchange and equity markets is shown to have a strong impact on the dynamic properties of high frequency returns. Therefore, only by taking account of this strong intraday periodicity is it possible to uncover the complex intraday volatility dynamics that exist both within and across different financial markets. This paves the way for new volatility modelling techniques that incorporate periodic components into the formulation of the conditional volatility equation.

One of the periodicity adjustment methods which has been proposed, and is gaining in popularity, is the periodic GARCH or PGARCH model developed by Bollerslev and Ghysels (1996). This technique makes it possible to explicitly incorporate periodicity into the parameters of any standard GARCH models through the conditional variance equation. Andersen and Bollerslev (1997), for example, applied a procedure based upon the flexible Fourier functional form (FFF) of Gallant (1981, 1982) to control and to account for the periodicity effects using the PGARCH structure. The application of this measure in the US foreign exchange market produced a marked improvement in the performance of GARCH models in producing forecasts. This finding is supported in a recent study in the UK futures market. Specifically, McMillan and Speight (2004b) report that applying the FFF method to standard GARCH models provides more consistent and reliable forecasting results.

As an alternative, Taylor (2004) introduces the spline version of the PGARCH model, which is capable of estimating different cubic spline functions between selected points (or knots) within a specific periodic cycle. This technique not only overcomes the rigidness of the functional form of the FFF-version of the PGARCH, but is also capable of producing superior model fit and forecasting performances. Further applications of these techniques on existing conditional volatility models have again documented their usefulness. This is evident in the superior model fit and forecasting performance produced when compared against the standard unadjusted models.[3] These recent developments augur well for the continuing usage of the GARCH models in volatility modelling and forecasting research.

Another recent issue that has become important in the return volatility literature concerns the measurement of the unobservable "true volatility". The most popular method used to measure *ex post* daily volatility is to use absolute demeaned daily returns or squared demeaned daily returns over the relevant forecasting horizons. However, Andersen and Bollerslev (1998a, 1998b) argue that the measure of "true volatility" based on *ex post* daily squared returns is problematic, as it includes a noisy component, which makes it an inefficient estimator. Moreover, Andersen and Bollerslev (1998a, 1998b) reason that the relative failure of the GARCH models arises not from a failure of the model but a failure to specify correctly the "true volatility" measure against which forecasting performance is compared.

As an alternative measure, Andersen and Bollerslev (1998a) introduced a new generation of conditional volatility models, which make use of a volatility measure

[3] Martens, Chang and Taylor (2002) find that the P-GARCH model provides the best forecasting performance and that the FFF-based variable is an efficient way of determining the periodic components of volatility. In addition, Taylor (2004) finds that the use of the spline-version of the PGARCH model produces more accurate forecasts and consistent VaR measures.

known as integrated realized volatility measure. This measure can be constructed by summing intraday squared returns. This allows the treatment of daily volatility as observed rather than latent, providing that the sampling frequency of squared returns is sufficiently high. By making use of the theory of quadratic variation and arbitrage-free processes, Andersen, Bollerslev, Diebold and Labys (2001, 2003) show that realized volatility constructed as above is not only model-free, but as the sampling frequency of the returns approaches infinity, has estimates that are measurement-error-free as well. However, as a cautionary note, a recent study has shown the potential of this measure in reviving the usefulness of the GARCH models in volatility modelling and forecasting. Specifically, McMillan and Speight (2004a) implement this technique in a study that comprises a dataset of 17 daily foreign exchange rate series. It is found that the GARCH model outperforms smoothing and moving average techniques, which have been previously identified as providing superior volatility forecasts.

1.1 Significance of the Study

As highlighted above, intraday periodicity has been widely observed in financial time series. Recent research examining intraday volatility dynamics in the developed markets reports that failure to account for this periodicity results in inconsistent GARCH parameter estimates in relation to the theoretical predictions on temporal aggregation. It has been observed that intraday volatility appears higher at the market opening and the market closing, resulting in a stylised U-shape pattern being reported for a variety of markets. Previous studies on the KLSE also indicate the presence of a U-shaped intraday volatility pattern in the return volatility (Mohammed, Fauzias and Othman, 1995, and De Brouwer, 2002). These studies, however, did not attempt to distinguish the trading periods into a morning session and an afternoon session. This could be an inaccurate assumption because studies of financial markets that are closed during the lunch hour indicate the presence of two distinct U-shaped patterns, i.e., one for the morning session and one for the afternoon session.[4] As the KLSE is also closed during the lunch hour, we will investigate whether a similar observation is valid for the KLSE using 5-minute returns obtained using KLCI data. This will lead to a better understanding of the dynamics of return volatility across the trading day for the KLSE.

The current study is also the first comprehensive attempt to compare the performance of three conditional volatility models within the parametric GARCH class of models using high frequency KLCI returns data. Previous studies on the Malaysian stock exchange relied on the symmetric GARCH and the GARCH–M models, with daily closing prices as the basis for the computation of return volatility (see Mohammed, Fauzias and Othman, 1995, and Chong, Ahmad and Abdullah, 1999) with lag parameters (p,q) of $(1,1)$ as the preferred specification. We also use lag structures (p,q) of $(1,1)$ to determine the best performing GARCH model but the returns are based on 5-minute returns data. In addition, this study utilises several recent methods developed to incorporate the effects of periodicity into intraday volatility modelling. We make use of a simplified version of the PGARCH processes introduced by Bollerslev and Ghysels (1996), which make it possible to explicitly incorporate the periodicity effects into the parameters of the GARCH models. In this aspect, we employ half-hourly and quarter-hourly dummy variables in the conditional volatility equation of the GARCH models. We also apply the FFF-based and the

[4] See, for example, Andersen, Bollerslev and Cai (2000), who found that intradaily volatility exhibits a double U-shaped pattern for the Japanese stock market based on a 4-year sample of 5-minute Nikkei 225 returns from 1994 to 1997. Taylor (2004) finds similar pattern for the cocoa futures market on the *Euronext.liffe* exchange based on 5-minute frequency returns on all futures contracts from 1997 to 2002.

spline-based variables suggested by Andersen and Bollerslev (1997, 1998a) and Taylor (2004), respectively, as the inputs into the PGARCH processes, to account for the intraday periodicity. As an alternative approach to control for the effects of periodicity, we also apply the two-step filtration method proposed by Andersen and Bollerslev (1997, 1998a) and Martens, Chang, and Taylor (2002). This latter approach differs in that there is a clear separation in the process of modelling the volatility and estimating the periodicity components. All these approaches are new modelling techniques that have not previously been applied to KLSE data.

To examine the relative quality of the various models employed, we also undertake an extensive analysis of the out-of-sample forecasts produced by the standard GARCH models and the various PGARCH specifications. To compare the predictive accuracy of alternative forecasts, we employ an asymptotic test of the null hypothesis of no difference in the accuracy of two competing forecasts proposed by Diebold and Mariano (1995). As an alternative, this study considers the test of forecast encompassing developed by Harvey, Leybourne and Newbold (1998). This aspect of forecast evaluation has been largely ignored or has never been applied to any study of volatility on the KLSE.

Another new contribution of this study to the literature is the use of GARCH forecasts in the construction of VaR models. VaR has become increasingly important in recent years as a measure of the market risk of a portfolio. The VaR models are widely used in financial and banking institutions as well as by market players. The adequacy and the quality of the VaR models developed in this study are evaluated using the three-step testing procedures proposed by Christoffersen (1998) and the regression-based tests of Clements and Taylor (2003). These are again newly introduced measures, which have not previously been applied to the KLSE.

Finally, we model the intraday dynamics of return volatility using the new integrated realized volatility measure introduced by Andersen and Bollerslev (1998a). We use the ARMA (1,1) specification to model the daily realized volatility with different intraday squared returns sampling frequencies. We then compare the forecasting performance of both the ARMA models and the GARCH models with the specially constructed daily realized volatilities as the proxy for the true daily volatility. We also compare the performances of both the ARMA and the GARCH models with the previous approach of using the demeaned squared returns as the proxy for the true daily volatility. We then construct VaR models based on both the most accurate ARMA and the GARCH forecasts, and ascertain whether these models provide adequate coverage through the various tests mentioned above. Again, all these applications are relatively innovative and we believe that this study is the first of its kind in the Malaysian context.

1.2 Justification of the Study

The application of new methods in the modelling and forecasting of KLCI returns is important because the KLSE is one of the largest emerging capital markets in Asia. It has a good chance of developing into a viable regional competitor to several currently large markets in the Asia Pacific region. It is important to recognise that the KLSE has different risk and return characteristics as well as different institutional structures from other developed markets. Many of the recently developed methods have not yet been evaluated in the context of the KLSE. Therefore, a major aim of this thesis is to examine whether the applications of these methods produce similar results to those reported for other markets. Certainly, this offers a window of opportunity to test and to evaluate the robustness of the various new measures, which have been found to be successful in more developed markets. The findings of the current study have important implications for those in risk-related industries, as well as for academics who are researching volatility in emerging and new financial markets. For example, the new modelling methods, which incorporate the effects of periodicity, could aid in the production of superior forecasts and, in turn, create better VaR models. The new methods could also be considered for other conditional volatility models (such as the stochastic volatility models), and the GARCH models used in this study could serve as a performance benchmark.

The introduction of the integrated realized volatility measure will help us to understand the true nature of volatility dynamics and offer an alternative to the demeaned squared returns volatility measure, which is widely used in the literature. The application of the integrated realized volatility measure will also provide a good opportunity to revaluate the relevancy and the adequacy of the GARCH models, which have been proven successful in the past. Finally, we hope to contribute to the scarce literature on emerging capital markets, considering that these markets form about 75% of the world's organised capital markets (Ariff, Shamsher and Annuar, 1998).

1.3 Objectives of the Study

The main focus of this study is to investigate the usefulness of the GARCH-based models in intraday volatility modelling and forecasting in the presence of periodicity. Specifically, this study has the following objectives:

1. To ascertain whether the two distinct U-shaped patterns, which are common to markets that are closed during the lunch hour, are observable for the KLSE.

2. To model the intraday volatility patterns using thirteen competing GARCH-based modelling approaches.
3. To assess the out-of-sample forecasting performances of the various GARCH models resulting from the applications of the thirteen modelling approaches.
4. To evaluate the quality of the forecasts generated in objective no.3, using both the Diebold and Mariano (1995) and the Harvey, Stephen and Newbold (1998) tests.
5. To evaluate the quality and adequacy of the various VaR models constructed from the GARCH forecasts and the *RiskMetrics* model, using both the three-step testing procedures proposed by Christoffersen (1998) and the regression-based tests of Clements and Taylor (2003).
6. To construct daily realized volatility measures based on the 1-minute, 5-minute, 10-minute, 15-minute and 30-minute intraday squared returns sampling frequencies as well as the one-day frequency realized volatility (equivalently, the daily squared returns).
7. To model the various daily realized volatility measures in objective no.6, using the ARMA (1,1) model.
8. To compare the forecasting performances of the ARMA models and the daily GARCH models using the various daily realized volatility measures as proxies for true volatility.
9. To evaluate the adequacy of the VaR models constructed from the ARMA, *RiskMetrics* and GARCH forecasts assessed in objective no.8, using both the Christoffersen (1998) and the Clements and Taylor (2003) tests.

1.4 Organisation of the Thesis

This study is divided into seven chapters. The present chapter has discussed the background and the rationale for the study, outlined the study's objectives, and briefly presented the methodology to be used. This chapter draws attention to the need for comprehensive research into volatility modelling and forecasting from a Malaysian perspective.

Chapter 2 provides an overview of the stylised facts regarding financial markets; these include volatility clustering, fat tails in the unconditional distributions of financial asset returns, mean reversion effects and leverage effects. It also describes in details the properties of the ARCH/GARCH processes. This chapter then provides some explanation for the U-shaped phenomena. In addition, it gives an overview of the PGARCH framework and the development of the integrated realized volatility measure, and provides some details of past findings on the performance of the GARCH models.

Chapter 3 gives an overview of the history and the development of the KLSE since its inception in the early 1960s. It then focuses on the performance of the KLCI and how the index is designed and computed. The chapter ends with a brief description of the trading practices of the KLSE.

Chapter 4 focuses on the intraday volatility pattern (i.e. periodicity) of returns across the trading day, taking into account the fact that the KLSE is closed during the lunch hour. It then discusses the methods and approaches used to model the intraday volatility periodicity on the KLSE. There are thirteen approaches used in this thesis, each based on one of the following models: the non-periodic (unadjusted) GARCH models, the jointly estimated GARCH models and the two-step filtration models. Both the jointly estimated GARCH and the two-step filtration models employ one of four variables in the conditional volatility equations. These variables are in the form of half hourly and quarter hourly dummy variables, the FFF-based variables and the spline-based variables. Augmented versions of the FFF-based and the spline-based variables using both the jointly estimated methods and the two-filtration techniques are also discussed.

Chapter 5 focuses on the forecasting performances of the thirteen volatility estimation approaches which utilise the PGARCH-based models and the non-periodic GARCH specifications described in Chapter 4. The performance of each approach is assessed using the mean squared forecast error (MSFE) and the mean absolute forecast error (MAFE). This chapter also discusses the applications of both the Diebold and Mariano (1995) and the Harvey, Leybourne and Newbold (1998) tests to compare the predictive accuracy of alternative forecasts. Finally, this chapter focuses on the evaluation of the adequacy of the VaR models produced by the available GARCH forecasts. For this purpose, both the Christoffersen (1998) and the regression-based tests of Clements and Taylor (2003) are applied.

Chapter 6 focuses on the integrated realized volatility measures. It describes how the various daily realized volatilise measures are designed and the ARMA models used to estimate them. The comparison of the out-of-sample forecasting performances of both the ARMA and the daily GARCH models using the various daily realized volatilities and the daily squared returns as proxies for the true daily volatility are then discussed. Finally, the chapter investigates the adequacy of the VaR models constructed from both the ARMA and GARCH forecasts. Again, both the Christoffersen (1998) and the regression based tests of Clements and Taylor (2003) are applied. The final chapter, Chapter 7, summarises the important conclusions based on the findings of the thesis.

CHAPTER 2

LITERATURE REVIEW

2.0 Introduction

The aim of this chapter is to present an overview of the literature on the areas covered in this thesis. The following discussion is divided into six sections. The first section begins with a discussion of the random walk hypothesis and the efficient market hypothesis. The second section touches on the stylised facts about the financial markets. Subsequently, issues on the properties of the ARCH and GARCH models, the U-Shaped pattern and the theories explaining this pattern, how volatility impacts financial markets, the performance of the GARCH model taking into account the periodicity factor, the availability of high frequency data and the integrated realized volatility measure will be discussed.

2.1 The Random Walk and the Efficient Market Hypotheses

The random walk hypothesis and the efficient market hypothesis are perhaps the earliest models proposed to explain the dynamics of financial assets. The random walk hypothesis asserts that financial asset price movements will not follow any patterns or trends and that past price movements cannot be used to predict future price movements. In the simplest terms, a random walk is a process whereby the previous change in the value of a variable is unrelated to future changes. In other words, a random walk defines the path of a random variable where each change or innovation is independent of all previous changes (implying zero correlation between successive pairs of observations) and each is drawn from an identical probability distribution (i.e. one with the same distributional parameters; the same mean and the same standard deviation). The efficient market hypothesis, on the other hand, encompasses broader concepts regarding the nature of asset prices. Despite its breadth, these concepts all imply one core feature: that financial markets are efficient transmitters of information that affects price. Much of the development of the theory on these subjects can be credited to the pioneering works of Bachelier (1900) and Cowles (1933, 1944), and subsequently, the works of Kendall (1953), Samuelson (1965), Fama (1965) and Roberts (1967). Fama (1970) assembles a comprehensive review of the theory and evidence of market efficiency. He concludes that the price of a stock reflects a balanced rational assessment of its true underlying value (i.e. rational expectations). This implies that the stock price at a particular time will have fully and accurately discounted (taken account of) all available information (news).

The efficient market hypothesis assumes several underlying conditions, which include perfect information, instantaneous receipt of news, and a marketplace with many small participants (rather than one or more large participants with the power to influence prices). The theory also assumes that news arises randomly in the future (otherwise the non-randomness would be analysed, forecast and incorporated within prices already). The theory predicts that the movements of stock prices will approximate stochastic processes, and that technical analysis and statistical forecasting will most likely be fruitless. Samuelson (1965), for example, eloquently summarises that a more efficient market will generate random price changes sequences, and the most efficient market of all is one in which price changes are completely random and unpredictable.

Therefore, in an efficient market, financial asset price movements can be described as

$$r_t = \frac{p_t - p_{t-1}}{p_{t-1}} = \mu_t + \varepsilon_t . \qquad \mathrm{E}[\varepsilon_t] = 0 , \; \mathrm{Var}[\varepsilon_t] = \sigma_t^2 . \tag{2.1}$$

where the return at time t, r_t, is the percentage change in the asset price p_t over the period from t-1 to t. This is equal to μ_t, a non-random mean return for period t, plus a zero mean random disturbance ε_t that is independent of all past and future $\varepsilon_t^{'}s$. It is the lack of serial correlation in the random $\varepsilon_t^{'}s$ that is the defining characteristic of efficient market pricing, i.e., past price movements reveals no information about the sign of the random component of the return in period t. Therefore, given this intuitive appeal, it is not difficult to see why the random walk hypothesis and the efficient market hypothesis have become icons of modern financial economics that continue to fire the imaginations of academics and professionals alike.[1]

However, several major works have challenged the supremacy of these two hypotheses. These works have managed to uncover empirical evidence which suggests that stock returns contained predictable components. Keim and Stambaugh (1986), for example, find statistically significant predictability in stock prices by using forecasts based on certain predetermined variables. In addition, Fama and French (1988) show that long holding-period returns are significantly negatively serially correlated, implying that 25 to 40 percent of the variation in longer horizon returns is predictable from past returns. Moreover, Lo and MacKinlay (1988) find that the random walk model is generally not consistent with the stochastic behaviour of weekly returns, especially for smaller capitalization stocks.

---

[1] For a brief and an excellent history of market efficiency, please refer to Dimson and Mussavian (1998).

At the same time, many researchers who have studied price movements in stock markets began to document stylized facts about financial markets - facts that present enough evidence to seriously challenge the validity of the random walk and the efficient market hypotheses.

2.2 Stylised Facts of Financial Markets

One of the earliest observations of these stylised facts concerns the volatility of financial asset returns: in particular, the phenomenon where certain periods are more volatile than others. Mandelbrot (1963) describes this market volatility as follows:

> "At closer inspection, however, one notes that large price changes are not isolated between periods of slow change...In other words, large changes tend to be followed by large changes - of either sign - and small changes tend to be followed by small changes..."
>
> (Mandelbrot, 1963, page 418)

This phenomenon later became known as *volatility clustering*, and is one of the many features of today's high frequency financial market data. Subsequent studies by Fama (1965), Chou (1988) and Schwert (1989) confirmed this observation. They also report that large changes in the price of an asset are often followed by other large changes, and small changes are often followed by other small changes. These observations are so universal that any casual observation of financial time series reveals bunching of high and low volatility patterns. In other words, volatility is positively correlated over time. The implication of such volatility clustering is that volatility shocks today will influence the expectation of volatility for many periods in the future.

Another stylised fact related to market volatility is what is known as *fat tails* in the unconditional distributions of financial asset returns. This phenomenon has been observed since the early 1960s, again by Mandelbrot (1963) and Fama (1963, 1965). They report that financial asset returns have leptokurtic unconditional distributions. Typical kurtosis estimates range from 4 to 50, indicating very extreme non-normality. The source of the heavy tails may be revealed by the relation between the conditional density of returns and the unconditional density. If the conditional density is Gaussian, then the unconditional density will have excess kurtosis if the conditional Gaussian densities have different variances. However, there is no reason to assume that the conditional density itself is Gaussian, and many studies assume that the conditional density is itself fat tailed, generating still greater kurtosis in the unconditional densities. Volatility clustering and the fat tails associated with financial asset returns are closely related. Indeed, while the latter is a static explanation, a key insight provided by volatility models such as the ARCH models is a formal link between dynamic (conditional) volatility behaviour and (unconditional) heavy tails. The ARCH models, introduced by Engle (1982) and the numerous extensions thereafter, as well as the stochastic volatility (SV) models, are essentially built to mimic volatility clustering.

Financial asset return volatilities also exhibit a property known as *mean reversion.* It is important to note that volatility clustering implies that volatility is temporal. Thus a period of high volatility will eventually give way to more normal volatility, and similarly, a period of low volatility will be followed by a rise in volatility. Mean reversion in volatility is generally interpreted as meaning that there is a normal level of volatility to which volatility will eventually return. Very long run forecasts of volatility should all converge to this same normal level of volatility, no matter when they are made. While most practitioners believe that this is a characteristic of volatility, they differ in view regarding the normal level of volatility and whether it is constant over time or not.

Another important stylised fact that is often quoted in the literature on market volatility is an observation known as the *leverage effect*. This term, coined by Black (1976), describes the tendency for volatility to rise more following a large price fall than following a price rise of the same magnitude. Black (1976) reasons that when the price of a company's stock falls, its value (of the equity) will also fall. As a result, the company's leverage or its debt-to-equity ratio will increase. Leverage is generally interpreted as an indicator of company riskiness. Therefore, when the leverage is high, the company is considered more risky, and a higher degree of risk or uncertainty entails higher volatility. Many proposed volatility models impose the assumption that the conditional volatility of the financial asset is affected symmetrically by positive and negative innovations. Clearly, this is inadequate because the presence of leverage effects implies an asymmetry in volatility clustering in financial markets. Therefore, there is a need for volatility models that can accommodate leverage effects, and this is one of the motivating factors lying behind many extensions of the basic ARCH models. Basic understanding of volatility asymmetry permits researchers to refine existing models by incorporating variables that account for the asymmetry in a more efficient way.

One of the objectives of this thesis is to model the intraday dynamics of index return volatility. We realize the importance of accommodating the stylised facts in the volatility models. Gigli (2002) explains the importance and the impact of the stylised facts on the design of a model. He reasons that empirical stylised facts often guide the specification of a model. A model's ability to reproduce such stylised facts is a desirable feature and failure to do so is most often a criterion to dismiss a specification. On the other hand, one does not try to explain all possible empirical regularities at once with a single model. Consequently, this thesis makes extensive use of the ARCH/GARCH models and their many extensions in modelling and forecasting return volatility.[2] Subsequent discussion will focus on the properties of these models and their applications, particularly with regard to modelling and forecasting return volatility.

2.3 The ARCH and the GARCH Models

Studies have indicated that many relationships in finance are intrinsically non-linear. For example, Campbell, Lo and MacKinlay (1997) assert that the payoffs to options are non-linear with respect to some of the input variables, and investors' willingness to trade off returns and risks is also non-linear. These observations provide clear motivations for consideration of non-linear models in a variety of circumstances in order to better capture the relevant features of the data. In short, much financial time series data exhibit features that could not be captured adequately by linear models. It can be said that a serious limitation of the linear models is their failure to account for changing volatility. For instance, the width of a forecast interval remains constant even as new data become available, unless the parameters of the model are changed.

[2] For a detailed survey of the various extensions of the basic ARCH/GARCH and stylised facts about return volatility, please refer to Bollerslev, Engle and Nelson (1994) and Hamilton (1994).

This is because linear models generally assume that the expected value of all error terms, when squared, is the same at any given point. This assumption is termed homoskedasticity. In contrast, data in which the variances of the error terms are not equal, in which the error terms may reasonably be expected to be larger for some points or ranges of the data than for others, are said to suffer from heteroskedasticity.

It has been shown that in the presence of heteroskedasticity, the regression coefficients from an ordinary least squares regression are still unbiased, but the standard errors and confidence intervals estimated by conventional procedures will be too narrow, giving a false sense of precision. Engle (2001) argues that instead of considering this as a problem to be corrected, ARCH and GARCH models treat heteroskedasticity as variance to be modelled. As a result, not only are the deficiencies of least squares corrected, but also a prediction is computed for the variance of each error term. The forecast intervals, therefore, are able to widen immediately to account for sudden changes in volatility, without changing the parameters of the model. Because of this feature, the ARCH and GARCH models have become valuable in the analysis of economic time series.

2.3.1 The ARCH Model

The ARCH model was introduced by Engle (1982). The acronym ARCH stands for AutoRegressive Conditional Heteroskedasticity. The term "heteroskedasticity" refers to changing volatility or variance. It must be stressed that it is not the variance itself which changes with time according to the ARCH model; rather, it is the *conditional* variance which changes, in a specific way, depending on the available data. The conditional variance quantifies our uncertainty about future observations, given all information available to date.

To provide a context for the ARCH model, let us consider the conditional aspects of a simple first-order autoregression equation

$$y_t = \alpha y_{t-1} + e_t \ . \tag{2.2}$$

where $t = \{t \in \mathbf{Z}^+ : 1 \le t \le T\}$, e_t is white noise with variance $V(u_t) = \sigma^2$, the unconditional mean of y_t is zero, while its conditional mean is αy_{t-1}. It has unconditional variance $V(y_t) = \sigma^2$, and its conditional variance is $V(y_t | y_{t-1}) = \sigma^2 / (1 - \alpha^2)$, so it is clear that the variance of this model is constant. Now consider the properties of the ARCH processes.

The proposed ARCH process consists of two equations. The first is the *conditional mean equation.* The simplest conditional mean equation could be based on the assumption that (log) returns, R_t, are generated under weak-form efficiency, as follows:

$$R_t = \mu + \varepsilon_t \qquad \varepsilon_t | \psi_{t-1} \sim D(0, h_t) \ . \tag{2.3}$$

where μ is the mean of the process, ψ_{t-1} is the set of information available at time t-1, and D is a Normal distribution with support over (-∞, ∞), a mean equal to zero and conditional variance equal to σ_t^2 or h_t, which is the volatility process to be estimated. The error term or the returns innovation process, ε_t, is then written as $\varepsilon_t = \sigma_t z_t$, with z_t an independent zero-mean, unit-variance stochastic process noise term. Equation (2.3) says that the *conditional distribution* of ε_t given ψ_{t-1} is normal, $D(0, h_t)$. In other words, given the available information ψ_{t-1} the next observation ε_t has a normal distribution with a conditional mean of $\mathrm{E}[\varepsilon_t|\psi_{t-1}] = 0$, and a conditional variance of $V[\varepsilon_t|\psi_{t-1}] = h_t$. We can think of these as the mean and variance of ε_t computed over all paths which agree with ψ_{t-1}.

The second equation in the ARCH process is the *conditional variance equation.* Suppose the ARCH (q) model for the series ε_t is defined by specifying the conditional distribution of ε_t, given the information available up to time t-1, and q is the lag parameter. The information set ψ_{t-1} is assumed to consist of all observed values of the series, and anything that can be computed from these values, e.g., innovations, squared observations, etc. We say that the process ε_t is ARCH (q) if the conditional distribution of ε_t, given the available information ψ_{t-1}, is

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 . \tag{2.4}$$

with $\alpha_0 > 0, \alpha_i \geq 0$ for all i, and $\sum_{i=1}^{q} \alpha_i < 1$.

Equation (2.4) specifies the way in which the conditional variance h_t is determined by the available information. Here h_t is not constant. Rather, it is made up of two components. The first is the constant term and the other is the dependence of the current variance on the size of the error term in the previous period. Notice that the conditional variance h_t is defined in terms of *squares* of past innovations. In other words, the volatility is modelled by allowing the conditional variance of the error term, h_t, to depend on the (immediately) prior value of the squared error. Therefore, if the error was large in the previous period (positive or negative), then the variance in period t will be higher. This, together with the assumptions that $\alpha_0 > 0$ and $\alpha_i \geq 0$, guarantees that h_t is positive, as it must be, since it is a conditional variance (a negative variance at any point in time would be meaningless).

The ARCH (q) model is nonlinear, since if ε_t could be expressed as $\varepsilon_t = \sum_{k=0}^{\infty} a_k e_{t-k}$, then we would have variance $V[\varepsilon_t | \psi_{t-1}] = V[\varepsilon_t | e_{t-1}, e_{t-2}, \ldots] = V[e_t]$, a constant. This contradicts equations (2.3) and (2.4), so $\{\varepsilon_t\}$ must be a nonlinear process. Moreover, since the model is nonlinear, the observations $\{\varepsilon_t\}$ in an ARCH (q) model are non-Gaussian. The distribution of $\{\varepsilon_t\}$ tends to be more fat-tailed than that implied by a normal distribution. Thus, outliers may occur relatively more often. This is a useful feature of the model, since it reflects the leptokurtic nature of returns observed in practice. Moreover, once an outlier does occur, it will increase the conditional volatility for some time thereafter. Once again, this reflects a pattern often found in real data. It may seem odd that, while the conditional distribution of ε_t given ψ_{t-1} is Gaussian, the unconditional distribution is not. The reason for this is that the unconditional distribution is an average of the conditional distributions for each possible path up to time t-1. Although each of these conditional distributions is Gaussian, the variances h_t are unequal. So we get a mixture of normal distributions with unequal variances, which is not Gaussian.

A practical problem encountered in fitting ARCH (q) models to financial asset returns data is that in order to obtain a good fit, the value of q, the number of lags of the squared error that are required to capture all of the dependence in the conditional variance, might be very large. Moreover, the non-negativity constraints described above might be violated if the value of q is very large (see Brooks, 2002). Everything else being equal, the more parameters there are in the conditional variance equation, the more likely it is that one or more of them will have negative estimated values.

2.3.2 The GARCH Model

An alternative and more flexible lag structure is provided by the Generalized ARCH, or GARCH (p,q) model proposed independently by Bollerslev (1986) and Taylor (1986). The full GARCH (p,q) model adds p autoregressive terms to the ARCH (q) specification. It gives a parsimonious representation that is easy to estimate and even in its simplest form, has proven successful in predicting conditional variances. The GARCH model allows the conditional variance to be dependent upon previous lags, so that the conditional variance equation in the simplest case is now

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 . \tag{2.5}$$

This is known as a GARCH (1,1) model. Again, h_t is the conditional variance, since it is a one-period ahead estimate for the variance based on any past information thought relevant. Using the GARCH model, it is possible to interpret the current fitted variance, h_t, as a weighted function of a long-term average value (dependent on α_0) of information about volatility observed during the previous period ($\alpha_1 \varepsilon_{t-1}^2$) and the fitted variance from the model obtained during the previous period ($\beta \sigma_{t-1}^2$).

To see how the GARCH model is more parsimonious than the ARCH model, consider the conditional variance equation of GARCH (1,1) above. Taking lags of the conditional variance equation in (2.5), the following expression would be obtained:

$$\sigma_{t-1}^2 = \alpha_0 + \alpha_1 \varepsilon_{t-2}^2 + \beta \sigma_{t-2}^2 . \tag{2.6}$$

Taking the lag of this equation gives,

$$\sigma_{t-2}^2 = \alpha_0 + \alpha_1 \varepsilon_{t-3}^2 + \beta \sigma_{t-3}^2 . \tag{2.7}$$

Substituting into (2.6) for σ_{t-1}^2

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta(\alpha_0 + \alpha_1 \varepsilon_{t-2}^2 + \beta \sigma_{t-2}^2), \tag{2.8}$$

$$= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_0 \beta + \alpha_1 \beta \varepsilon_{t-2}^2 + \beta^2 \sigma_{t-2}^2 . \tag{2.9}$$

Now substituting into (2.9) for σ_{t-2}^2

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_0 \beta + \alpha_1 \beta \varepsilon_{t-2}^2 + \beta^2 (\alpha_0 + \alpha_1 \varepsilon_{t-3}^2 + \beta \sigma_{t-3}^2), \tag{2.10}$$

$$= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_0 \beta + \alpha_1 \beta \varepsilon_{t-2}^2 + \alpha_0 \beta^2 + \alpha_1 \beta^2 \varepsilon_{t-3}^2 + \beta^3 \sigma_{t-3}^2, \tag{2.11}$$

$$= \alpha_0 (1 + \beta + \beta^2) + \alpha_1 \varepsilon_{t-1}^2 (1 + \beta L + \beta^2 L^2) + \beta^3 \sigma_{t-3}^2, \tag{2.12}$$

where L is the lag operator. An infinite number of successive substitutions of this kind would result in

$$\sigma_t^2 = \alpha_0 (1 + \beta + \beta^2 + \ldots) + \alpha_1 \varepsilon_{t-1}^2 (1 + \beta L + \beta^2 L^2 + \ldots) + \beta^\infty \sigma_0^2 . \tag{2.13}$$

The first expression on the right-hand side of (2.13) is a simple constant and, as the number of observations tends to infinity, β^∞ will tend to zero. Hence the GARCH (1,1) model can be written as

$$\sigma_t^2 = \gamma_0 + \alpha_1 \varepsilon_{t-1}^2 (1 + \beta L + \beta^2 L^2 + \ldots), \tag{2.14}$$

$$= \gamma_0 + \gamma_1 \varepsilon_{t-1}^2 + \gamma_2 \varepsilon_{t-2}^2 + \gamma_3 \varepsilon_{t-3}^2 + \ldots \tag{2.15}$$

Equation (2.15) is a restricted infinite order ARCH model. Thus the GARCH (1,1) model, containing only three parameters in the conditional variance equation, is a very parsimonious model that allows an infinite number of past squared errors to influence the current conditional variance.

The GARCH (1,1) model can be extended to a GARCH (p,q) formulation, where the current conditional variance is parameterized to depend upon q lags of the squared error and p lags of the conditional variance,

$$\alpha_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \ldots + \alpha_q \varepsilon_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2 + \ldots + \beta_p \sigma_{t-p}^2, \tag{2.16}$$

or

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}. \tag{2.17}$$

Nelson and Cao (1992), and Drost and Nijman (1993) give the necessary and sufficient conditions to ensure non-negativity of conditional variance in (2.17), and the process is covariance stationary if and only if $\alpha_1 + \alpha_2 + \ldots + \alpha_q + \beta_1 + \beta_2 + \ldots + \beta_p < 1$. To demonstrate the latter condition, consider again the following GARCH (1,1) model:

$$R_t = \mu + \varepsilon_t. \qquad \varepsilon_t \big| \psi_{t-1} \sim D(0, \sigma_t^2) \tag{2.18}$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta h_{t-1}. \tag{2.19}$$

where $\alpha_0 > 0, \alpha_1 \geq 0, \beta \geq 0$. If $\alpha_1 + \beta < 1$, then the unconditional variance of ε_t is given by

$$V[\varepsilon_t] = \alpha_0 / (1 - \alpha_1 - \beta) \tag{2.20}$$

For $\alpha_1 + \beta \geq 1$, the unconditional variance of ε_t is not defined and hence is non-stationary.

In many applications, especially with daily frequency financial data, the estimate for $\alpha_1 + \alpha_2 + \alpha_3 + \ldots + \alpha_q + \beta_1 + \beta_2 + \beta_3 + \ldots + \beta_p$ turns out to be very close to unity. The sizes of the parameters α and β determine the short-run dynamics of the resulting volatility time series. Large β coefficients indicate that shocks to conditional variance take a long time to die out, so volatility is "persistent". Large α coefficients mean that volatility reacts quite intensely to market movements, and so if α is relatively high and β is relatively low, then volatilities tend to be more "spiky".

The introduction of the GARCH model spurred a vigorous line of research leading to a number of variants of the GARCH (p,q) model. Many of the extensions of the GARCH model have been suggested as a consequence of perceived problems with the standard GARCH (p,q) model. Basically, there are three inadequacies that need to be addressed; they are:

1. The non-negativity conditions that might be violated by the estimated model. The only way to overcome this is to place artificial constraints on the model coefficients in order to force them to be non-negative.
2. The GARCH models cannot account for the leverage effects described earlier. The leverage effect stems from the fact that the standard GARCH model enforces a symmetric response of volatility to positive and negative shocks. This arises since the conditional variance equation is a function of the magnitudes of the lagged residuals and not their signs. In other words, by squaring the lagged error in the conditional variance equation, the sign is lost.
3. The standard GARCH model does not allow for any direct feedback between the conditional variance and the conditional mean.

This thesis will not attempt to investigate or discuss all the possible variants of the standard GARCH (p,q) model. Instead, the reader is invited to refer to the excellent survey conducted on the subject by Bollerslev, Chou and Kroner (1992). What this thesis will focus on are four different variants, which have lifted some of the restrictions and inadequacies of the basic GARCH model. The models in question are the Exponential GARCH (EGARCH) model proposed by Nelson (1991), the Threshold GARCH (TGARCH) model of Zakoian (1994) and the GARCH in Mean (GARCH–M) models suggested by Engle, Lilien and Robins (1987) through the ARCH-M specification. A brief description of each of these models is given in Chapter 4. We now turn our discussion to the U-shaped pattern in the intraday volatility of asset returns.

2.4 The U-Shaped Pattern

Empirical research on equity markets, using stock market transactions data, has revealed several intraday regularities in the returns on stock prices. Perhaps the most interesting regularity is the U-shaped intraday pattern in the volatility of asset returns and in the volume of trading in markets with an overnight close. The existence of a U-shaped pattern in volatility across the trading day could be dated back to at least the findings of Wood, McInish and Ord (1985) and Harris (1986) in the US markets. Harris (1986), for example, examined the phenomena on the New York Stock Exchange (NYSE) and found that prices rise sharply during the first 45 minutes of trading and that returns are high near the very end of the day, particularly on the last trade of the day. Furthermore, it is observed that the day-end price changes are greatest when the final transaction is within the last five minutes of trading.

Regarding trading volume, Jain and Joh (1988) studied hourly aggregate NYSE volume and found that the volume is particularly high at the beginning and towards the close of trading. Similar studies by Brock and Kleidon (1992) and Foster and

Viswanathan (1993) yielded U-shaped patterns in both volume and volatility of returns in NYSE stocks. These studies exhibited a significant positive relationship between the volume and volatility of the stocks; for example, the highest volume coincides with the highest variance, which incidentally is more frequent at both the open and the close than for the rest of the day. Studies in other markets find similar results: Yadav and Pope (1992) for the UK market, Choe and Shin (1993) for the Korean market, and Lam and Tong (1999) for the Hong Kong market. The existence of intraday patterns and, more specifically, U-shaped patterns are not exclusive to equity markets. These patterns have been demonstrated for other markets as well. These include the findings of Peterson (1990) and Aggarwal and Gruca (1993) for the equity options markets; Baillie and Bollerslev (1991) for the foreign exchange markets; Cornett, Schwarz and Szakmary (1995) for the foreign exchange futures markets; and Jordan, Seale, Dinehart and Kenyon (1988) and Taylor (2004) for the commodities markets.

In addition, studies in markets that have a break during the lunch hour indicate the presence of a double U-shape pattern instead of a single one. Andersen, Bollerslev and Cai (2000), in a study of the Nikkei 225 on the Tokyo Stock Exchange using 5-minute frequency returns, find that the intraday return volatility exhibits a double U-shaped pattern associated with the opening and closing of the separate morning and afternoon sessions. They report that Nikkei 225 index volatility is significantly higher at the opening of the morning and the close of the afternoon sessions than during the mid-morning and mid-afternoon sessions. These features, combined with an increased volatility immediately before and after the lunch break, result in two distinct U-shapes: one in the morning and one in the afternoon. Taylor (2004), in a study using 5-minute frequency returns of cocoa futures contracts on the *Euronext.liffe* exchange, also reports a similar finding. It is found that return volatility is high at the opening of trading, both at the beginning of the trading day and at the beginning of trading after the lunchtime period. This pattern is repeated at the close of trading, both at the end of the trading day and just before the lunchtime period.

Summarising, it is crucial to recognize and identify what kind of impact and influence the intraday patterns may have on modelling return volatility. Many recent studies have indicated that studies that utilize daily, weekly and monthly prices will not be able to explain the intraday dynamics of return volatility (see Andersen and Bollerslev, 1998a). The availability of high frequency data has made it possible to do this and many of the recent studies in the ARCH/GARCH literature, for example, have employed a higher frequency, typically at the 5-minute and 15-minute intervals, to model return volatility. This presents a unique opportunity to test the robustness of the GARCH model and its many extensions. We now discuss some of the theories that attempt to explain the dynamics of the U-shaped pattern.

2.5 Explanations for the U-Shaped Pattern

The existence of the U-shaped patterns in the volatility of asset returns and in the volume of trading has generated a strong interest in finding the appropriate models to explain and understand the origin of these phenomena. The three models that are most often quoted are the asymmetric information hypothesis of Admati and Pfleiderer (1988) and Foster and Viswanathan (1990) and the increased demand hypothesis of Brock and Kleidon (1992). All these models consistently assume that the existence of heterogeneous investors, combined with periodic market closure, results in discretion by investors in timing their trades, which can lead to an endogenous concentration of trades and price changes. Admati and Pfleiderer (1988) argued that the interaction between informed and uninformed discretionary investors gives rise to the U-shaped pattern. Uninformed discretionary investors are defined as those investors who have the ability to choose when to trade during the day on the basis of trading costs. It is reasonable to assume that there are times during the day that attract both informed and uninformed discretionary investors to trade, which in turn result in a concentration of trade volume. The reason why informed investors will operate in the market at a certain time is because there will also be uninformed investors through whom they can disguise their trades. Similarly, uninformed discretionary investors will choose to trade when they perceive that there are increased activities of informed investors relative to other periods. These are the times when the trading costs are at the optimal level. The interaction of both parties at these times results in a clustering of volume, which could explain concentrations of volume such as the one just before the market closes or at market openings. The model cannot, however, predict with certainty whether a concentration of trading volume should occur at the opening of the market, in the middle of the day or at the close.

Foster and Viswanathan (1990) propose that the intraday pattern is due to the interaction between informed traders and a subset of the discretionary liquidity traders who act strategically. The informed traders receive information each day, but the value of this information diminishes through time because there is a public announcement of some of the private information each day. In this way, discretionary liquidity traders have an incentive to delay their transactions when they believe that the informed traders are particularly well-informed. By waiting, they can learn from the trades that occur and the public signal that is released. On the other hand, the informed traders, knowing that there is a forthcoming public signal, trade more aggressively, and thus more information is released through that trading. The delay tactic by the discretionary liquidity traders leaves less liquidity in the market and makes it easier for the market makers to infer the informed traders' reasons for trading. As a consequence, the volume is lower and prices are more volatile.

Brock and Kleidon (1992) argue that most trading in the beginning and at the end of the trading day results from the inability to trade when the market is closed. It is argued that the risk of holding positions overnight when trade is not possible differs from that of holding them during the day when continuous trading strategies can be followed. This may cause traders to adjust their positions at the end of the day to account for the change in trading possibilities. The end of the day is therefore likely to be a period when the volume of non-discretionary liquidity trading is high. Symmetrically, discretionary liquidity traders will especially trade in the morning at the opening of the following day, due to the accumulation of overnight information in the absence of an opportunity to trade. Brock and Kleidon (1992) also observe that there is a tendency for passive portfolio managers to trade at the end of the day. This is a consequence of the fact that the performance of these funds is based on how closely the fund tracks a specified index. Because the value of the index is calculated using closing prices, passive portfolio managers can reduce the tracking error by trading at the end of the day. For example, they can decide to sell (to obtain a paying investment) at the end of the day when the index is decreasing to do better relative to the index. Another reason for trade clustering at the end of the day may be due to the rising industry of day traders who close positions before exchanges close so as not to have overnight positions. This course of action is more likely to happen at the end of the day.

Gerety and Mulherin (1992) go further by focusing on the assumption that investors differ in their willingness and/or ability to hold positions overnight. Accordingly, if these investors transfer the risk of holding a position while the market is closed, then the volume at the end of the day should be directly related to the volatility expected to occur overnight. Correspondingly, the trading activity at the opening is positively related to unexpected overnight volatility. On the other hand, Kim and Verrecchia (1994) point out that an anticipated information event may affect not only the trading pattern during the transition between trading and non-trading periods, but also during other periods in the trading day. This takes place before the announcement event if the anticipated public announcement stimulates private information-gathering and trading. When the announcement is released, investors form posterior beliefs and trade on their private information and market prices. After the announcement, given slow dissemination of earnings information, excess portfolio rebalancing activities may result during the trading day. We will discuss below some of the findings that made use of high frequency data and the impact on the adequacy of the ARCH and GARCH models in modelling and forecasting return volatility.

2.6 Review of Past Findings

2.6.1 Volatility and the Financial Markets

The temporal behaviour of stock market volatility has fascinated many researchers since the 1970s. Malkiel (1979) and Pindyck (1984) contend that the upward trend in volatility during the last thirty to forty years in the US is the reason for the decline in equity prices. Pindyck (1984) argues that the variance of the firms' real gross marginal return on capital has increased significantly since 1965. This has increased the relative riskiness of investors' net real returns from holding stocks. He believes that this can explain the reason for the market decline. To a certain extent, this argument is in line with Black's (1976) findings that stock returns tend to be strongly negatively correlated with changes in volatility.

Porteba and Summers (1986) investigate this issue by testing the time-series properties of volatility. It is argued that shocks to volatility have to persist for a very long time in order for volatility to have a significant impact on stock prices. If shocks to volatility are only transitory, no adjustment of the future discount rate will be made by the market. Therefore, expected stock returns are not affected by the volatility movement. This finding contradicts the claims of Malkiel (1979) and Pindyck (1984) above. Pindyck (1986), however, responded by estimating a portfolio choice model. It is reasoned that even though the changes in variance do not persist for long, they do provide a better explanation for the market decline compared to other variables such as changes in corporate profits and changes in the real interest rate. In fact, about one-third of the 1974 market decline can be attributed to volatility changes.

Chou (1988) supports the findings of Pindyck (1986), using the GARCH estimation technique to study volatility persistence. The findings highlight the weaknesses of the two-stage OLS estimation methodology employed by Porteba and Summers (1986). Specifically, it is argued that this method is inadequate and less efficient than the maximum-likelihood methods used in GARCH estimation. It is demonstrated that the GARCH (1,1) – M model, in particular, is a more suitable tool than the two-stage methods due to the quality of results obtained. It is also demonstrated that the parameter estimates using the two-stage methods are very sensitive to the frequency of volatility measurements and that the monthly measure used by Porteba and Summers (1986) tends to seriously underestimate the persistence parameter. Furthermore, it is argued that the assumption of constant conditional means and variances is both unrealistic and inaccurate. Subsequently, it is discovered that the persistence of shocks to the stock return volatility is so high that the data cannot distinguish whether the volatility process is stationary or not. Assuming stationarity, it is found that the half-life of the volatility shocks is about one year. The parameter estimates and the non-stationary test results are both robust to changes in the frequency of data measurements. The results confirm the findings of Malkiel (1979) and Pindyck (1984) above and at the same time emphasize the need to accommodate the effect of heteroskedasticity when estimating return volatility.

2.6.2 Performance of the GARCH Model

There is a vast literature that attempts to compare the performance and accuracy of the GARCH model with other volatility models in producing out-of-sample volatility forecasts. Among the earliest research to extensively test the properties of the ARCH/GARCH processes was the study carried out by Akgiray (1989), who finds evidence that time series of daily stock returns exhibit significant levels of dependence that cannot be modelled as a linear white-noise process. A reasonable return-generating process is empirically shown to be a first-order autoregressive process with conditionally heteroskedastic innovations. It is found that the GARCH (1,1) model in particular fits the data very satisfactorily. A comparison of forecasts of 24 monthly return variances using four methods - the simple historical average, the exponentially weighted moving average (EWMA), the ARCH and the GARCH models - concluded that the ARCH and GARCH models could simulate the actual pattern of stock market volatility more closely than the simple historical average and EWMA methods and that the GARCH specification is superior to the ARCH. The results of the simple historical average forecasts do not reflect short-term changes in volatility and are virtually unchanged throughout the 24-month period. The same is also true for the EWMA forecasts, which are found to inadequately model the transitory changes in volatility. These findings show that the time-series behaviour of market volatility can be realistically modelled by conditionally heteroskedastic processes.

The analysis proceeds by evaluating the various model-based forecasts using a number of forecast error statistics. Based on the relative values of these statistics, it is found that the GARCH forecasts are far better than the other three, and that this is even more so in periods of high volatility. GARCH forecasts are also generally less biased, as evinced by the smaller values of the forecast error statistics obtained. Therefore, it is concluded that the GARCH forecasts constitute substantial improvement over the traditional forecasts such as the historical sample averages.

In a similar study, Pagan and Schwert (1990) compare the EGARCH, GARCH, Markov switching regime, and three non-parametric models for forecasting monthly US stock return volatilities. The results indicated that for the US stock market from 1834 to 1937, the EGARCH is the best volatility forecasting model. The GARCH models performed moderately while the remaining models produce very poor forecasts. Finally, West and Cho (1995) compare the out-of-sample performance of univariate homoskedastic, GARCH, autoregressive, and non-parametric models for conditional variances, using five bilateral weekly exchange rates for the dollar. It is found that for a one-week horizon, GARCH models tend to produce slightly more accurate forecasts. However, for longer horizons, it is difficult to find grounds for choosing between the various models.

While the modelling and forecasting of US stock market conditional volatility has found some support for the GARCH framework, the analysis of conditional volatility in other international stock markets has produced conflicting results. Tse (1991) examines stock return volatility in the Tokyo Stock Exchange. Based on fitted ARCH and GARCH models in the period from 1986 through 1987, the forecasts of return volatility in 1988 through 1989 are produced. The ARCH/GARCH forecasts are compared with a benchmark value, a naive forecast and an exponential weighted moving average (EWMA) forecast. The results show that the EWMA method produces the best forecasts. Tse and Tung (1992) apply the same tests in the Singapore stock market and obtain the same results.

In a more extensive study, Franses and van Dijk (1996) compare the volatility forecasting performance of the GARCH, quadratic-GARCH (QGARCH, Engle and Ng, 1993), and TGARCH models in comparison to the random walk model using weekly German, Dutch, Italian, Spanish and Sweden stock index returns over the period from 1986 to 1994. It is found that the random walk model performs particularly well when the crash of 1987 is included in the estimation sample, while the QGARCH model performs better upon its exclusion.

Brailsford and Faff (1996) examine the ability of various volatility models to forecast aggregate monthly stock market volatility in Australia. The models tested include the random walk, historical mean, moving average, exponential smoothing, EWMA, a simple regression model, two standard GARCH models and two asymmetric Glosten, Jagannathan and Runkle (GJR) GARCH models. Using several loss functions, they are unable to identify a clearly superior model and suggest that the best forecasting model depends upon the subsequent application. The rankings of the various model forecasts are sensitive to the choice of error statistic. However, the study finds some support for the GARCH models, particularly the GJR-GARCH (1,1) specification, which is consistently ranked high, together with the simple regression model.

McMillan, Speight and Gwilym (2000) analyse the predictive power of several classes of GARCH models (Standard GARCH, TGARCH, EGARCH, Component-GARCH), against the random walk, historical mean, moving average, exponential smoothing, EWMA and regression models, for the FTSE 100 (FTSE) and FT All Share (FTA) Indices in the UK. The comparison was carried out using data obtained over monthly, weekly and daily frequencies, using both symmetric and asymmetric loss functions in the evaluation of forecasts. Under symmetric loss, the results suggest that the random walk model provides vastly superior monthly volatility forecasts, while the random walk, moving average and recursive smoothing models provide moderately superior weekly volatility forecasts, and the GARCH, moving average and exponential smoothing models provide marginally superior daily volatility forecasts. When asymmetric loss is considered, it is found that the ranking of forecasting methods is dependent on the series, frequency and the direction of the asymmetry. The historical mean shows the best result for the forecasting of daily FTA and FTSE volatility. The historical mean and simple regression are jointly favoured for weekly FTA volatility, and exponential smoothing is favoured for forecasting weekly FTSE volatility. However, if attention is restricted to one forecasting method for all frequencies, the most consistent forecasting performance is provided by the moving average and GARCH models.

There are very few studies published on the performance of the GARCH models in emerging stock markets. Gokcan (2000) compares the forecasting performance of the GARCH (1,1) model against the EGARCH (1,1) model using monthly returns from seven emerging stock markets, including Argentina, Brazil, Colombia, Malaysia, Mexico, the Philippines and Taiwan. The results indicate that for all the countries except Brazil, the GARCH (1,1) model produces superior model fit and out-of-sample forecasting performance. Similarly, Chong, Ahmad and Abdullah (1999) study the performance of six variations of GARCH models against the random walk model using five daily observed Malaysian stock market indices (Composite Index, Tins Index, Plantations Index, Properties Index, and Finance Index) on the KLSE. The results indicate that the EGARCH (1,1) model is the best model in forecasting volatility for all five stock market indices. This is followed by the symmetric GARCH (1,1) model and the non-negative GARCH (1,1) model. The unconstrained GARCH (1,1) and the GARCH–M (1,1) models are both ranked fourth. The random walk model is ranked second last while the Integrated GARCH (1,1) model is ranked last.

Poon and Granger (2003) carried out an extensive survey of the volatility forecasting research over the last twenty years. They focus on four types of volatility forecasting methods that have been widely used in the literature; viz. historical volatility

(HISVOL) models, ARCH/GARCH models, SV models and option-implied volatility (ISD) based on the Black-Scholes model. Poon and Granger (2003) define the scope of the historical volatility method such that it includes the random walk and historical averages of squared returns or absolute returns. Also included are moving averages, exponential weights, autoregressive models, fractionally integrated autoregressive absolute returns and the more sophisticated techniques like the multivariate realized volatility model in Andersen, Bollerslev, Diebold and Labys (2003). In order to find the method that will produce the best forecasts, they carried out pair-wise comparisons involving 66 studies that include the four volatility forecasting methods. Interestingly, for studies involving both HISVOL and ARCH/GARCH models (39 related studies in total), twenty-two found HISVOL better at forecasting than ARCH/GARCH, and seventeen found ARCH/GARCH superior to HISVOL. When they examined eighteen studies involving both ARCH/GARCH and ISD models, they found that seventeen studies are in favour of the ISD compared to only one in favour of the ARCH/GARCH models. In another comparison, which involves four studies on ARCH/GARCH and SV models, three are in favour of the SV models as opposed to only one for the ARCH/GARCH models. Poon and Granger (2003) also find that in studies involving ARCH/GARCH models, the GARCH models produce better forecasts than the ARCH models. In general, models that incorporate volatility asymmetry, such as the EGARCH and GJR-GARCH models, perform better than the standard GARCH models.

The overall ranking suggests that ISD provides the best forecasts, followed by HISVOL and ARCH/GARCH with roughly equal performance. The superior performance of ISD is expected because the forecasts are based on a larger and timelier information set. Poon and Granger (2003) explain that the options markets are small compared to the equity markets and in most emerging markets, for instance, options are not traded at all. Therefore, time series models will continue to influence the direction of volatility forecasting even though they are inferior to the ISD models. They also highlight the possible problem of bias in the publication of these studies, i.e., papers presented are prepared for different reasons, use different data sets, many kinds of assets, various interval frequencies, a variety of evaluation techniques, and face the pressure of conforming to support a viewpoint for a particular method in the publication process.

2.6.3 High-Frequency Data and the Periodicity Factor

The increased availability of high-frequency financial data has spurred research interest in the complex nature of intraday-return dynamics. The applicability of the standard GARCH model as an adequate description of volatility in intraday financial data has been called into question recently. In particular, studies examining intraday foreign exchange rate and index futures data have reported GARCH coefficients that are inconsistent with those reported at the daily level in the light of theoretical results on the temporal aggregation of GARCH processes (see Andersen and Bollerslev, 1997). The problem is made worse when it is also documented that the standard GARCH models are also incapable of modelling satisfactorily the return volatility, which varies systematically over the trading day.

Andersen and Bollerslev (1997) demonstrate the problem of direct ARCH modelling of intraday return volatility in the presence of pronounced systematic fluctuations in the return series. It is argued that standard ARCH models imply a geometric decay in the return autocorrelation structure and simply cannot accommodate strong regular cyclical patterns (intraday periodicity), which have a strong impact on the autocorrelation patterns of the 5-minute return series employed in their study. Instead, the combination of recurring cycles at the daily frequency and a slow decay in the average autocorrelations may be explained by the joint presence of the pronounced intraday periodicity (the U-shaped pattern for example) coupled with the strong daily conditional heteroskedasticity, which can be modelled sufficiently with standard ARCH models. They suggest that high frequency volatility modelling should start on this premise, i.e., awareness of the interaction between the interdaily conditional heteroskedasticity and the intraday periodicity. Similarly, McMillan and Speight (2004a) reason that the importance of identifying an appropriate method of periodicity adjustment and reliable GARCH model estimation follows directly from the fact that the relative frequency of intraday observations, compared with identifiable shocks, is much greater than that afforded by interday data. It is found that estimates of parameters are only consistent when the periodicity effects are taken into account.

Several periodicity adjustment methods have been introduced in recent years to control for the periodicity effect in intraday volatility modelling. The more popular ones are the methods introduced by Bollerslev and Ghysels (1996) and Andersen and Bollerslev (1997, 1998a). Bollerslev and Ghysels (1996) introduced the periodic GARCH (PGARCH) framework, which is designed to capture the repetitive periodic time variation in the second-order moments. It is claimed that the PGARCH model provides a natural generalization of the time-invariant seasonal GARCH models to allow for a greater degree of flexibility when modelling periodicity in the conditional variances. The PGARCH framework includes all GARCH models in which a set of periodic intercept dummy variables is included in the variance equation.

Meanwhile, Andersen and Bollerslev (1997, 1998a) advocate a procedure based upon the FFF variables, which employs a series of trigonometric terms to identify the systematic component of the return series. It is claimed that this method would aid in uncovering the complex link between the short and long run return components, which in turn may help to explain the apparent conflict between the long memory volatility characteristics observed in interday data and the rapid short run decay associated with news arrivals in intraday data.

Martens, Chang and Taylor (2002), using a GARCH (1,1) model for the original returns as the benchmark, show that modelling the intraday seasonal (periodicity) volatility pattern improves the out-of-sample volatility forecasting. In addition to the FFF variables, which are used in conjunction with the PGARCH structure, they also introduce a two-step approach in modelling the intraday periodicity in the foreign exchange market. First, they estimate the seasonal component by applying an OLS regression of squared returns (or absolute returns) on either the FFF variables or the dummy variables constructed under the PGARCH framework. The standardized return series from this regression is then considered as the filtered or deseasonalized return series. In the second step, they estimate the parameters of the GARCH model based on the filtered return series. The results indicate that the PGARCH model provides the best forecasting performance, followed by the two-step filtration approach, and lastly by the standard GARCH model without any adjustment for periodicity.

The advantage of controlling for the periodicity in intraday return series is also highlighted by Taylor (2004) using UK commodity futures data. The PGARCH model once again produces superior forecasts of future return volatility compared to other competing volatility models and the standard GARCH model. Similarly, McMillan and Speight (2004a), using data from the stock index futures market in the UK, report that GARCH models that use returns that are adjusted for periodicity (using the FFF method) provide better model fit and produce superior forecasting results when compared with similar models that utilize unadjusted returns data. The subject of periodicity adjustment will be dealt with in greater detail in Chapter 4 of this thesis.

2.6.4 The Integrated Realized Volatility Measure

Another recent development that has revived the usefulness of the GARCH model is the introduction of a new volatility measure by Andersen and Bollerslev (1998a). In this paper, they argue that the failure of the GARCH model to provide good forecasts, which is reported throughout the literature, is not a failure due to the properties of the GARCH model itself, but a failure to specify correctly the "true volatility" measure against which forecasting performance is measured. Andersen and Bollerslev (1998a) reason that the standard approach of using *ex post* daily squared returns as the measure of "true volatility" for daily forecasts is flawed because this measure includes a large and noisy zero mean constant variance error term, which is unrelated to actual volatility. Let us consider again equation (2.1) above. A common approach for judging the forecast performance of any model is to compare its predictions with subsequent realizations. Since volatility is not a directly observable process, this approach is not immediately applicable. However, if the model for σ_t^2 is correctly specified, then $\mathrm{E}_{t-1}(R_t^2) = \mathrm{E}_{t-1}(\sigma_t^2 z_t^2) = \sigma_t^2$, which appears to justify the use of squared returns innovation over the relevant horizon as a proxy for *ex post* volatility. However, while the squared innovation provides an unbiased estimate of the volatility process, it may yield very noisy measurements due to the idiosyncratic term, z_t. This component typically displays a large degree of observation-by-observation variation relative to σ_t^2, such that the proportion of variability in squared returns that can be attributed to volatility is low. This is the reason why volatility models often report poor predictive power. Consequently, an alternative measure for "true volatility" is suggested based upon the cumulative squared returns from intraday data. This measure, which is referred to as integrated realized volatility, allows more meaningful and accurate volatility forecast evaluation. Subsequently, it is found that the forecasting performance of a GARCH (1,1) model is improved when the daily volatility is measured by means of the cumulative squared intraday returns. It is also demonstrated that the variance is substantially smaller the higher the frequency used to generate the integrated realized volatility. Therefore, with the availability of high frequency returns, the *ex post* realized daily volatility should be measured using the highest frequency.

Recent research in this area has provided evidence to support the superiority of the new volatility measure when compared to the squared returns measure used previously. An example of this is the findings of a study conducted by Martens (2001). Multiple period forecasts from intraday volatility models are compared with forecasts from daily volatility models. When these forecasts are evaluated using integrated realized volatility, the results show that the higher the frequency used, the better the daily volatility forecast from the relevant GARCH (1,1) model becomes. This forecast is simply constructed from multiple out-of-sample forecasts for frequencies higher than the daily frequency. The GARCH (1,1) model for intraday returns also gives better forecasts than augmenting the daily GARCH (1,1) model to include the difference between the daily high and low. Extending the daily model with the sum of squared intraday returns leads to a similar performance as modelling the intraday returns directly.

In a recent study, McMillan and Speight (2004a) reconsider the accuracy of the GARCH-based volatility forecasts compared to those produced by exponential smoothing and moving average models for seventeen daily exchange rates relative to the US dollar in the foreign exchange market. The measure of "true volatility" used to evaluate forecasts is based upon 30-minute intraday observations and the models were estimated over a five-year in-sample period with a one-year out-of-sample forecasting period. The results show that the GARCH models outperform both the exponential and the moving average models for sixteen out of the seventeen currencies in the sample. This is almost the complete reverse of results that previously showed that the GARCH models consistently underperform compared to the two statistical averaging models.

Very few studies on integrated realized volatility have been conducted in emerging capital markets. One recent study is conducted on the Indian Stock Exchange. Pandey (2003) compares the empirical performance of various unconditional volatility estimators and conditional volatility models (GARCH and EGARCH) using time-series data on the S&PCNX Nifty, a value-weighted index of 50 stocks traded on the National Stock Exchange (NSE), Mumbai. The estimates computed by various estimators and conditional volatility models over non-overlapping one-day, five-day and one-month periods are compared with the "realized volatility" measured over the same period. The data set used to construct measures of realized volatility is based on three years' (1999-2001) five-minute frequency returns. In order to test the ability of the estimators and models to forecast volatility, the estimates of unconditional estimators are compared with the realized volatility measured in the next period of the same length. For conditional volatility models, the forecasts for the same periods are obtained by estimating models from the time-series prior to the forecast period. The results indicate that while conditional volatility models provide less biased estimates, extreme-value estimators are more efficient estimators of realized volatility. As far as the forecasting ability of the models and estimators is concerned, conditional volatility models fare extremely poorly in forecasting the five-day (weekly) or monthly realized volatility. In contrast, extreme-value estimators generally perform relatively well in forecasting volatility over these horizons.

In summary, this chapter has highlighted the finer points of the random walk hypothesis and the efficient market hypothesis in relation to assets pricing and returns volatility. It has also discussed the stylised facts about financial markets. The properties of the ARCH and GARCH models are elaborated in detail, followed by a discussion of the intraday U-Shaped pattern, which is a common observation in financial markets. The theories behind the occurrence of the U-shaped pattern are also discussed. The chapter ends with a discussion of the impact of volatility on the financial markets and the impact of periodicity and high frequency data in the forecasting performance of the GARCH model. The final part of this chapter also discusses the integrated realized volatility measure, why it is better than the squared returns measure and how it is able to revive the usefulness of the GARCH model. The following chapter, Chapter 3, looks at the history and the evolution of the KLSE. It also discusses how the KLCI is designed and computed. The KLCI returns data will be used extensively in this thesis. Chapter 3 ends with a brief explanation of the trading practices on the KLSE.

CHAPTER 3

THE KUALA LUMPUR STOCK EXCHANGE

3.0 Introduction

The aim of this chapter is to provide some information about the KLSE. The following discussion is divided into five sections. The first section gives an overview of the history and the development of the KLSE. The second section touches on company listings on the KLSE. Subsequent discussions focus on the KLCI, the performance of the KLCI, and finally on how stock trading is conducted on the KLSE.

3.1 The History and Development of the KLSE

In Malaysia, the KLSE[1] is the only stock exchange approved by the Minister of Finance under the provisions of the Securities Industry Act, 1983. The KLSE is a self-regulatory organization with its own memorandum and articles of association, as well as rules which govern the conduct of its members in securities dealings. The KLSE is also responsible for the surveillance of the marketplace and for the enforcement of its listing requirements, which spell out the ten criteria for listing, disclosure requirements and standards to be maintained by listed companies.

Although the history of the KLSE can be traced to the 1930s, the public trading of shares in Malaysia only really began in 1960, when the Malayan Stock Exchange (MSE) was formed. When the Federation of Malaysia was formed in 1963, with Singapore as a component state, the MSE was renamed the Stock Exchange of Malaysia (SEM). With

[1] The KLSE changed its name to the Bursa Malaysia Berhad following the successful demutualization of the Malaysian Securities Commission and the KLSE in April 2004.

the secession of Singapore from the Federation of Malaysia in 1965, the common stock exchange continued to function, but as the Stock Exchange of Malaysia and Singapore (SEMS). The year 1973 was a major turning point in the development of the local securities industry, for it saw the split of SEMS into the Kuala Lumpur Stock Exchange Berhad (KLSEB) and the Stock Exchange of Singapore (SES). The split was opportune in view of the termination of the currency interchange ability arrangements between Malaysia and Singapore. Although the KLSEB and SES were deemed to be separate exchanges, all the companies previously listed on the SEMS continued to be listed on both exchanges. When the Securities Industry Act 1973 was brought into force in 1976, a new company called the Kuala Lumpur Stock Exchange (KLSE) took over the operations of the KLSEB as the stock exchange in Malaysia. Its function was to provide a central marketplace for buyers and sellers to transact business in shares, bonds and various other securities in Malaysian listed companies. On 1 January 1990, following the decision on the "final split" of the KLSE and SES, all Singaporean incorporated companies were de-listed from the KLSE and vice-versa for Malaysian companies listed on the SES. The KLSE became a public company limited by shares, as opposed to its previous status as a company limited by guarantee, in January 2004. Subsequently, in April 2004, the KLSE officially launched its new name, Bursa Malaysia, together with a new organization structure. The holding company is now known as Bursa Malaysia Berhad. In March 2005, Bursa Malaysia Berhad made its debut on the Main Board. Bursa Malaysia's market capitalization as of 5 February 2005 stands at over RM700 billion (over USD 173 billion).[2]

[2] Source: Commerce International Merchant Bankers Berhad (2005).

3.2 Company Listing on the KLSE

Since the inception of the KLSE in the 1960s, the Main Board was the only board for company listing until the establishment of the Second Board in November 1988. The Second Board complements the Main Board and provides an opportunity for smaller firms that have great potential to grow but do not meet the listing requirements of the Main Board. Each board is further classified by sectors, which reflect the core business of these companies. As part of its aggressive push to become an Asian-Pacific hub for information and communication technology (ICT), in October 1997, the Government of Malaysia launched the Malaysian Exchange of Securities Dealing and Automated Quotation (MESDAQ) as a third board of the Malaysian stock market. The MESDAQ, which is modelled on the NASDAQ, is intended as an avenue for small and medium enterprises in technology-related areas to raise capital in order to establish a base in the multimedia super-corridor south of Kuala Lumpur. To give an example of the current state of affairs, as of 6 June 2006, a total of 648 companies were listed on the Main Board and 259 companies were listed on the KLSE Second Board, as well as 118 companies listed on the MESDAQ Market, giving a total of 1025 total companies listed on the KLSE. This is an increase of almost five-fold since the inception of the exchange back in 1973. At that time, the number of listed companies was only 262.

3.3 Kuala Lumpur Composite Index

The KLSE computes an index for each of the main sectors traded on the bourse, and currently there are 14 indices used to indicate the performance of each of the various economic sectors. However, the most widely followed, by far, is the Kuala Lumpur

Composite Index (KLCI). The KLCI[3] was introduced in 1986 after it was found that there was a need for a stock market index that would serve as a more accurate indicator of the performance of the Malaysian stock market and the economy. The KLCI is a weighted index, which was introduced in 1986 but extended back to January 1977, with 1977 as the base year. Prior to 1986 there was effectively no index that represented the entire market. The main indices used then were the KLSE Industrial Index (an all-shares value-weighted index), the New Straits Times Industrial Index (a 30-share price-weighted index) and the OCBC Composite Index (a 55-stock multi-sector based value weighted index). The rapid growth of the Malaysian economy saw the need for an encompassing index that could reflect the relationship between the market and the economy. The impressive growth and increased economic performance of KLSE listed companies meant that the three indices above were not adequate in absorbing and reflecting these rapid changes. The KLCI, introduced in 1986, was therefore designed to overcome these limitations. The following objectives were sought:

1. It should effectively reflect the performance of the companies listed on the stock exchange;
2. It should be generally sensitive to the investors' expectations;
3. It should be generally indicative of the impact of government policy changes;
4. It should be reasonably responsive to the underlying structural changes in the different sectors of the economy.

It started with a base of 67 component stocks in its augural year, 86 stocks until April 1995 and 100 stocks thereafter.

[3] Previously the KLCI was known as the KLSE Composite Index. The change of name was officially established in February 2005.

The KLCI design carefully takes into account the composition of the component stocks included in the computation of the index, in that they must reflect the sectoral developments of the economy. To ensure that the component stocks do not over or under-present certain sectors, the number of component stocks selected for different economic activities is also constantly correlated with the sectoral contribution to gross domestic product. The KLCI is computed by the market capitalization of each component stock as the weight, and the arithmetic mean as the method of averaging. Thus, the sum over all component stocks of the truncated mean of the daily closing prices, P_0, of a component stock in 1977 multiplied by the number of shares outstanding, Q_0, on 1 January 1977 is used as the Opening Base or Base Aggregate market value (Base AMV). The index on the first trading day of 1977 (3 January 1977) is given by

$$Index = \frac{\sum_{1}^{n} P_1 Q_1}{\sum_{1}^{n} P_0 Q_0} \tag{3.1}$$

where $\sum P_1 Q_1$ is the Current AMV and $\sum P_0 Q_0$ is the Base AMV. n is the number of component stocks, and P_1 and Q_1 are the daily closing price of a component stock and the corresponding number of shares outstanding as of 3 January 1977, respectively.

The following formulas are used to adjust the aggregate market value for rights issues, and inclusion and exclusion of a component stock into the index:

1. Rights Issue

   Adjusted Base AMV =

$$Old\ Base\ AMV \times \frac{Old.Current.AMV + Market.Value.of.Rights.Issue}{Old.Current.AMV}$$

2. Inclusion of a Component Stock

   Adjusted Base AMV =

$$\text{\textit{Old Base AMV}} \quad \text{x} \quad \frac{Old.Current.AMV + Market.Value.of.Included.Component.Stock}{Old.Current.AMV}$$

3. Exclusion of a Component Stock

   Adjusted Base AMV =

$$\text{\textit{Old Base AMV}} \quad \text{x} \quad \frac{Old.Current.AMV - Market.Value.of.Included.Component.Stock}{Old.Current.AMV}$$

The old current AMV is the aggregate market value of all component stocks based on the closing prices on the last day of cum-rights or on the last day before the inclusion or exclusion of a component stock. Similarly, the market value for rights, the market value of included component stocks and the market value of excluded component stocks are calculated on the same basis. No adjustment is made for bonus issues or stock splits, as there is no change in the aggregate market value. Since 3 January 1977, there were numerous occasions when adjustments were made for rights issues, inclusion and exclusion of component stocks. Thus, the KLCI is constantly updated to take any such changes into account.

3.4 The Performance of the KLCI

In terms of performance, the KLCI reached its highest peak at 1275.32 points at the end of 1993. This was an increase of 98% over the level at the end of 1992. At this time, it was ranked third among the world's top performers. The KLCI outperformed the indices of several developed and regional bourses, including the Tokyo (2.9%), New York (13.7%), London (20.1%), Singapore (59%) and Bangkok (88.4%) markets, but was lower than the gains of 115.7% and 154.4% in the Hong Kong and Manila markets, respectively. The KLCI performances remained bullish thereon, and it managed to remain above the 1000-point barrier until the occurrence of the Asian Currency Crisis in 1997. From the high of 1271.57 points on 25 February 1997, the KLCI declined by 794 points or 62.5% to 477.16 points on 12 January 1998. Due to the further declining health of the corporate sector and the high level of non-performing loans shouldered by the banking sector, the KLCI suffered its lowest level ever at 262.70 points on 1 September 1998. Nevertheless, prices rebounded strongly from 2 September 1998 to 7 September 1998 as the KLCI gained 69% to close at 445.06 points. The periods under study (2001 and 2002) saw the KLCI hover between 600 and 800 points, clearly on the path to recovery. Since then, the KLCI has continued on its upward path. As of 26 May 2006, the KLCI stood at 930.75 points.

3.5 Trading on the KLSE

In the early days, trading on the KLSE was conducted through an open-outcry system, where stock and share prices were determined through the bid and ask levels shouted out by traders on the trading floor of the KLSE. Since 1992, the KLSE has operated a fully automated trading system. All buy-ins and odd lots trading are fully automated. All trades are executed via the 'SCORE' or System on Computerised Order Routing and Execution maintained by the KLSE. Dealers will now key the bid and ask prices into their Broker Front End (BFR) terminals, which will electronically transmit all orders to SCORE. SCORE will then match these prices and orders. Once matching is completed, SCORE will confirm the successful transactions back to the dealers and also channel the same data to the Securities Clearing Automated Network System Sendirian Berhad or 'SCANS' for clearance. SCANS, which is a subsidiary of the KLSE, will be responsible as a co-ordinator to clear all trades transacted between the brokers. In other words, all payments for and delivery of stocks and shares are made by the brokers, on behalf of their clients, to SCANS and vice versa. Today, such functions are also fully automated via desktop banking and the central depository system.

Prior to July 1992, physical scrips that represented the stocks and shares issued by public listed companies were widely used for delivery in settlement of trades. To enhance settlement efficiency, the scripless system or Central Depository System (CDS) was fully implemented in November 1992. Under the CDS, all stocks and shares issued and traded on the KLSE will merely be book entries into and out of investors' securities or CDS accounts maintained with the Malaysian Central Depository Sendirian Berhad or MCD. The role of the MCD (a subsidiary of the KLSE) is to maintain a fully computerised Register of Depositors and administer the book entries for movement of stocks and shares transacted from one investor's CDS account to another.

Trading takes place five days a week (Monday-Friday), except on public holidays and other market holidays (when the Exchange is declared closed by the Bursa Malaysia Committee). There are two trading sessions on any market day: the morning session from 9:00 a.m. to 12:30 p.m. and the afternoon session from 2:30 p.m. to 5:00 p.m. The transaction day is denoted as day "t", whilst the following trading day (a day when KLSE is open) will be denoted as day "t+1" and so forth. Orders may be entered between 8:00

a.m. and 12:30 p.m. and between 2:00 p.m. and 5:00 p.m. The KLSE orders entered for each of the two trading sessions in a day are good for that session only. Unexecuted orders at the end of a trading session have to be re-entered into the system for execution.

Trades transacted on the KLSE must be cleared via the t+3 Rolling Settlement System (T+3 RSS), whereby settlement must conclude no later than t+3, i.e., the third trading day after the transaction day. There are three types of settlement basis under the T+3 RSS:

1. Ready Basis

   Payments for purchases transacted on day t must be made no later than 12.30 p.m. on day t+3. Delivery for a sale contracted must be made no later than 12.30 p.m. on day t+2. Hence, a seller has to ensure that there is sufficient credit balance of the stocks and shares that he has sold in his CDS account before they are due for delivery.

2. Designated Basis

   Stocks and shares that have been declared "designated counters" by the KLSE will follow the designated basis settlement period. This means that a seller must have sufficient stocks and shares in his CDS account prior to placing a sell order with his broker. Likewise, a buyer of designated stocks will need to make payment for the stocks and shares prior to placing a purchase order with his broker.

3. Immediate Basis

   A seller will have to ensure that the stocks and shares are available for sale in his CDS account not later than 12.30 p.m. on day t+1, whereas buyers must make payment to their brokers not later than 12.30 p.m. on day t+2.

To summarise, this chapter has provided an overview of the history and the development of the KLSE. It has also discussed the types of board that are available for company listing on the KLSE. The chapter then discusses why and how the KLCI was established and designed. Finally, the chapter discusses the current trading practices on the KLSE. The next chapter, Chapter 4, will focus on the investigation of the intraday volatility dynamics of the KLCI using 5-minute frequency returns data. The same set of data will also be extensively used in Chapter 5. The next chapter will introduce several modelling approaches mainly based on the PGARCH models in conjunction with the jointly estimated and the two-step filtration estimation techniques.

CHAPTER 4

INVESTIGATING INTRADAY VOLATILITY DYNAMICS

4.0 Introduction

The work presented in this chapter is the first of the three major investigations concerning the dynamics of intraday volatility on the KLSE. The second investigation focuses on the evaluation of performance and quality of various volatility forecasts produced by competing modelling approaches that employ GARCH-based models. We also evaluate the adequacy of the various VaR models constructed from the available volatility forecasts. All these works will be discussed in detail in Chapter 5. The third and final investigation centres on the modelling and forecasting of daily realized volatility. We will ascertain whether the adoption of the daily realized volatility as a proxy for the true daily volatility will improve the forecasting performance of the standard daily GARCH-based models. This work will be discussed in Chapter 6.

In this chapter, we attempt to establish the existence of the double U-Shaped periodicity pattern for the Malaysian market. We also investigate the usefulness of the GARCH-based models in modelling this intraday periodicity pattern using high frequency data. In this respect, we make comprehensive performance comparisons between thirteen competing modelling approaches. In order to ascertain the potential of the PGARCH models, we formulate twelve out of the thirteen modelling approaches using the PGARCH structure, while the remaining modelling approach is the standard unadjusted GARCH-based model. The best performance is determined by assessing which modelling approach has the best model fit.

This chapter is divided into four major sections. In the first section, we discuss the background of the work that will be covered in the chapter. In the second section, we provide details on the properties of the various conditional volatility models as well as the descriptions of the various modelling approaches. In the third section, we provide details on the data and the test procedures, as well as the results. We finish in the fourth section with a summary of the major findings. All results are presented at the end of the chapter.

4.1 Chapter Background

Estimates of asset return volatility are used to assess the risk of many financial instruments. Extensive research has shown us that volatility is the single most important variable in finance and it has become a vital component for consideration in investment management, security valuation, risk management and hedging strategies. Moreover, with the rapid growth in volatility-dependent financial derivative markets and products, the need for more sophisticated methods of measuring volatility becomes more crucial.

Another significant application of asset return volatility forecasts is in the application of VaR models. Manganelli and Engle (2001), for example, define VaR as the maximum potential loss in the value of a portfolio of financial instruments for a given probability over a certain horizon. Today, VaR has become the standard measure that financial analysts use to quantify market risk (market risk estimates the uncertainty of future earnings, due to changes in market conditions). The application of VaR has become so widespread that currently central banks in many major money centres, led by the Basle Committee on Banking Supervision (Basle Committee, 1996), require their supervised banks to measure the market risk of their assets and trading books within a VaR framework.

Given the above requirements, understanding how to obtain reliable measures (and forecasts) of asset volatility and how their dynamics evolve over time is essential. Research in this field is very active and many leading papers initially relied on the daily data in producing volatility estimates and forecasts. The increasing availability of financial market data at intraday frequencies has resulted in a change in focus. Moreover, it has led to the development of better *ex post* volatility measurements as well as an important information source for volatility forecasts. To this end, it is widely observed that return volatility varies systematically over the trading day, and this pattern is highly correlated with the intraday variation of trading volume and bid-ask spreads.

Many empirical studies have shown that standard time series models are inefficient with regards to modelling the dynamics of the intraday return volatility process.[1] In particular, new time series models need to take account of the seasonal or periodic volatility patterns that most high-frequency asset returns exhibit. These models should also be able to deal with well-known characteristics which are common to many financial time series. These include volatility clustering, i.e. the tendency of large absolute changes to be followed by large absolute changes and small absolute changes tend to be followed by small absolute changes; leptokurtosis (fat-tailed ness) in the unconditional distribution of financial time series returns; and the "leverage effect",

[1] See, for example, Andersen and Bollerslev (1998a), who attribute this to the inadequacy of the standard time series models of volatility when applied to high frequency returns data. Their analysis of intraday volatility patterns in the DM-USD foreign exchange and S&P 500 equity markets demonstrated that traditional time series methods, when applied to raw high frequency returns, may give rise to erroneous inferences about the return volatility dynamics.

which refers to the negative correlation between changes in stock prices and volatility. The introduction of conditional volatility models has made it possible to capture these characteristics, and in certain cases, to explicitly model the intraday periodicity patterns.

The introduction of the periodic GARCH (PGARCH) models by Bollerslev and Ghysels (1996) made it possible to explicitly incorporate periodicity into the parameters of the model. They show how practical estimation and extraction of the intraday periodic component of return volatility is both feasible and indispensable for a meaningful intraday dynamic analysis. Particular attention is focused on the differing impact of the periodic pattern on the dynamic return features at the various intraday frequencies. This is a significant development because it demonstrates that not only could the PGARCH model high frequency financial data more effectively than previous GARCH-based models, but also that it could successfully model periodically the systematic patterns in average volatility across the trading day.

To this end, many empirical studies using high frequency intraday data from a variety of markets indicate that PGARCH models give superior return volatility forecasts than those produced from standard GARCH models.[2] Taylor (2004) points out that many of the PGARCH modelling applications thus far have used data that are characterized by a U-shaped intraday volatility pattern and it may not be appropriate to use the existing PGARCH models if the volatility pattern is otherwise characterized. In order to overcome this deficiency, Taylor (2004) introduces augmented versions of the PGARCH models that allow for more complex conditional volatility dynamics, i.e., models capable of

[2] See, for example, Martens, Chang and Taylor (2002), for an application using the DM/USD and the YEN/USD exchange rates; Clements and Taylor (2003), for an application using FTSE100 index futures data; and Taylor (2004), for an application using data on the cocoa futures on LIFFE.

allowing for intraday return volatility patterns that may not conform to the U-shaped pattern.

Although the PGARCH model is potentially more efficient than the standard GARCH-based models, it poses a problem in that a large number of coefficients are required if there are many time periods included within each periodic cycle under study. In this respect, the model could become less parsimonious and the computation time involved may result in difficulty in estimating the periodic conditional return volatility. One solution to this problem is to apply the FFF-based variables advocated by Andersen and Bollerslev (1997, 1998a) in conjunction with the PGARCH model. The FFF version of the PGARCH model proves to be parsimonious and, more importantly, allows for smooth volatility dynamics. Recent studies by Martens, Chang and Taylor (2002) and Taylor (2004) indicate that this approach provides a highly significant improvement over the use of standard GARCH models in forecasting future return volatility. However, Taylor (2004) argues that the FFF version of the PGARCH model is rather restrictive because the technique assumes equality in conditional volatility at the beginning and end of the periodic cycle (due to the patterns generated by the cosine and sine functions). As an alternative, Taylor (2004) introduces the spline version of the PGARCH model, which is capable of estimating different cubic spline functions between selected points (or knots) within a specific periodic cycle. This technique not only overcomes the rigidness of the functional form of the FFF version of the PGARCH, but is also capable of producing superior and consistent VaR measures.

An interesting alternative to the simultaneous models of conditional volatility and periodicity described above is the two-step filtration approach proposed by Andersen and

Bollerslev (1997, 1998a) and employed by Martens, Chang and Taylor (2002). The two-step filtration approach is attractive because it is computationally less expensive than the jointly estimated PGARCH models that include periodic components. The first stage of the technique involves estimating and extracting the seasonal pattern, i.e. removing the periodicity from the financial data. The seasonal pattern could be estimated by either using simple intraday means of (log) squared returns or the fitted values from an ordinary least squares regression of (log) squared returns on FFF-based variables. The second stage is to estimate the filtered or adjusted data using GARCH-based models. A recent study by Martens, Chang, and Taylor (2002) indicates that modelling using the two-step approach based on FFF variables performs only marginally worse than similarly defined jointly estimated PGARCH models. Further modelling using the two-step filtration approach based on different financial markets with different microstructures could well be important in determining the robustness and the potential of this technique.

The KLSE, which is the focus of this study, is different from other established markets in that the trading session closes over the lunch time period i.e., it has two trading sessions. Studies of markets that are closed during lunch hours indicate that asset return volatility follows a double U-shaped pattern over the trading day.[3] Previous studies on the U-shaped intraday volatility pattern of asset return on the KLSE include the works of Mohammed *et al.* (1995) and De Brouwer (2002). Neither of these studies, however, attempted to distinguish the trading periods into a morning session and an afternoon session. This results in a single U-shaped pattern in volatility of returns across the trading day, i.e., following the dynamics of trading for markets that have a single trading session

[3] See, for example, Chang *et al.* (1993), and Andersen, Bollerslev, and Cai (2000), for the Japanese market; Cheung *et al.* (1994) for the Hong Kong market; and Bildik (2000) for the Turkish market.

instead of two in a trading day. Taylor (2004) argues that for markets that are closed during lunch hours, the intraday returns should not be characterized by a smooth U-shaped pattern because during the lunch time period, accumulated information may be compounded into prices at the opening of the afternoon session, resulting in an abrupt and discontinuous increase in return volatility at this point of time. It is to this gap in the literature that this chapter contributes. The results using high frequency KLCI returns data show that as in many other Asian markets which close during lunch hours, the KLSE does exhibit a double U-shaped intraday periodicity in return volatility.

This chapter also compares the performance of several different conditional volatility models within the parametric GARCH class of models on high frequency KLCI returns data. The following specifications of GARCH models were analysed: GARCH (generalized ARCH), EGARCH (exponential GARCH) and TGARCH (threshold GARCH). In order to evaluate whether these models adequately capture the volatility process and the intraday pattern of return volatility, we employ the periodic versions of these models (generically referred to as PGARCH models) introduced by Bollerslev and Ghysels (1996). We compare the results of volatility modelling using four competing variables incorporated into the conditional volatility equation of the five GARCH-based models above. The four variables used in the estimations are:

1. Half-hourly dummy variables equally spaced throughout the trading day,
2. Quarter-hourly dummy variables positioned at the opening and the closing of the trading period and quarter-hourly dummy variables positioned just before and after the lunch time period,
3. Flexible Fourier form based variables,

4. Spline variables based on selected points within the trading day.

This estimation approach is known as the joint estimation technique of the PGARCH model. Based on the same four variables, we estimate the parameters of the three GARCH models using the two-step filtration technique of Andersen and Bollerslev (1997, 1998a). The testing methodology is described in detail in Section 4.2. We use the non-periodic GARCH models as the benchmark to evaluate the performance of PGARCH models estimated jointly and estimated using the two-step technique. The aim is to ascertain whether the joint estimation and two-step filtration techniques offer significant advantages or contributions in terms of superior model fits over the standard GARCH models.

It is believed that this is the first study of its kind on the KLSE. The contribution of this study is that it not only provides an assessment of new techniques in modelling the intraday periodicity of the KLSE, but the modelling techniques also utilize high frequency 5-minute returns KLCI data that has not been employed in any of the earlier studies. We hope this effort will lead to a better understanding of the intraday volatility dynamics of the index returns in this market. We believe with better modelling techniques, we could improve on the accuracy and the quality of forecasts of the stock index. This is crucial for asset pricing and hedging, considering that the KLCI is also used as the underlying basis for stock index futures trading.

4.2 Conditional Volatility Models

This section gives a brief description of the non-periodic and periodic conditional volatility models adopted in the study. Throughout this description, the models considered will reflect the fact that KLCI returns are used in the empirical section.

4.2.1 Non-Periodic GARCH-based models

All the GARCH class of models used in this study consists of a linear mean and volatility equation. The mean equation is based on the assumption that (log) returns, R_t, are generated under weak-form efficiency, thus

$$R_t = \mu + \varepsilon_t \qquad \varepsilon_t | \psi_{t-1} \sim N(0, \sigma_t^2) \tag{4.1}$$

where $t = \{t \in \mathbf{Z}^+ : 1 \le t \le T\}$, μ is the mean of the process, ψ_{t-1} is the information set available at time t-1, and N is a continuous distribution with support over (-∞, ∞) and mean equal to zero and conditional variance equal to σ_t^2 (also denoted as h_t). With the exception of the GARCH–M models, we will use this particular mean equation for the rest of the GARCH models applied in this study, i.e., the GARCH models will differentiate themselves by changes in the specification of the volatility equation.

We note that information arrival in financial markets is clustered (hence conditional variance, h_t, is time-dependent). Therefore, the volatility equations of GARCH models analysed in the study are formulated such that current conditional variance is parameterised to depend upon q lags of the squared error and p lags of the conditional variance.

4.2.1.1 GARCH (*p,q*) model

This model assumes that conditional variance, h_t, is a weighted average of past squared residuals with weights that approach zero. The GARCH model also allows the conditional variance to be dependent upon previous lag.

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} \, . \tag{4.2}$$

The process is covariance stationary if and only if $\alpha_1 + ... + \alpha_q + \beta_1 + ... + \beta_p < 1$.

4.2.1.2 EGARCH (*p,q*) model

The Exponential GARCH (EGARCH) model introduced by Nelson (1991) assumes that h_t is an asymmetric function of past ε_t's as defined by:

$$\ln h_t = \omega + \sum_{i=1}^{q} \alpha_i g(z_{t-i}) + \sum_{j=1}^{p} \beta_i \ln(\sigma_{t-j}^2) \, . \tag{4.3}$$

where $z_t = \varepsilon_t / \sigma_t$ is the normalized residual series. The value of $g(z_t)$ depends on several elements. Nelson (1991) suggests that to accommodate the asymmetric relation between stock returns and volatility changes, the value of $g(z_t)$ must be a function of both the magnitude and the sign of z_t.

4.2.1.3 TGARCH (*p,q*) model

This model was introduced by Rabemananjara and Zakoian (1993) and is able to capture asymmetric responses to positive and negative errors in the conditional variance,

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \gamma \varepsilon_{t-1}^2 d_{t-1} + \sum_{j=1}^{p} \beta_j h_{t-j} . \qquad (4.4)$$

To allow for asymmetry in volatility, the basic GARCH (p,q) model is augmented by including a dummy variable, d_{t-1}, that takes the value of unity if $\varepsilon_{t-1}<0$, and zero otherwise. Asymmetry in volatility is inferred if $\gamma \neq 0$ and a leverage effect is present in the data if the estimated value of γ is positive.

4.2.2 Jointly Estimated Periodic GARCH model

In order to incorporate the periodic variation in any standard GARCH model, Bollerslev and Ghysels (1996) propose the inclusion of a set of periodic dummy variables in the conditional volatility equation of the particular GARCH model. For example, the PGARCH model of Bollerslev and Ghysels (1996) allows all coefficients in (4.2) to take a different value for each s time period within the periodic cycle of length S, where $s = \{s \in \mathbf{Z}^+ : 1 \leq s \leq S\}$ and s and t are related by a function denoted by $s(t)$ such that

$$(s,t) \in \{(1,1), (2,2), ..., (S-1, S-1), (S,S), (1, S+1), (2, S+2), ..., (S,T)\}.$$

However, for reasons of parsimony, a restricted version of their model is considered in this study. In particular, only the constant term in (4.2) in the conditional equation is allowed to vary over the periodic cycle. Under this assumption, the standard PGARCH model can now be expressed as follows:

$$h_t = \boldsymbol{\omega}' \mathbf{D}_t' + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} , \qquad (4.5)$$

where

$$\boldsymbol{\omega}' = \lfloor \omega, \tilde{\omega}' \rfloor$$

$$\tilde{\boldsymbol{\omega}}' = [\omega_2',\ldots,\omega_s',\ldots,\omega_S'],$$

$$\mathbf{D}_t' = [1, \tilde{\mathbf{D}}_t'],$$

$$\tilde{\mathbf{D}}_t' = [D_{2,t},\ldots,D_{s,t},\ldots,D_{S,t}],$$

and the dummy variable, $D_{s,t}$, takes a value of unity if the current observation is in the sth stage of the periodic cycle, and a value of zero otherwise. The dummy variables could also be similarly formulated for other conditional volatility models in (4.2.1.1), (4.2.1.2) and (4.2.1.3). In this study, this method is known as *the jointly estimated dummy version of the PGARCH model.*

Taylor (2004) warns that a potential problem with the above model is that a large number of coefficients may be required if the number of periodic dummy variables incorporated is large, i.e., if there are many time periods included within each periodic cycle (i.e., S is large). This makes the approach more expensive in terms of the time needed to estimate the parameters of the models. Taylor (2004) suggests that it might be possible to sidestep this problem by selecting periodic dummy variables that span more than one time period. However, this assumes that conditional volatility is constant within the time period covered by the dummy variables and then changes abruptly whenever a new time period is entered.

Andersen and Bollerslev (1997, 1998a) provide a solution to this problem and propose the use of the FFF-based variables to model periodic conditional volatility. As mentioned previously, this form can be used in conjunction with a PGARCH model and if we apply this to (4.2), for example, then the formulation will be as follows:

$$h_t = \sum_{q=1}^{Q}\left(\delta_{c,q}.\cos\left(\frac{2\pi qs(t)}{S}\right) + \delta_{s,q}.\sin\left(\frac{2\pi qs(t)}{S}\right)\right) + \sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2 + \sum_{j=1}^{p}\beta_j h_{t-j}, \tag{4.6}$$

Where S is the number of return intervals per day, Q is the tuning parameter to determine the order of the Fourier expansion, and δ is the coefficient of the FFF-based variable. In this study, this approach is referred to as *the jointly estimated FFF version of the PGARCH model.*

Again, a similar formulation can be used in conjunction with the other non-periodic volatility models described above. Taylor (2004) argues that this approach may not be adequate for markets that are closed over the lunch period due to the disruption in the continuity of conditional return volatility. The problem lies with the measure of time used with the periodic components estimated in the conditional volatility equation. Taylor (2004) argues that the periodic components in (4.6) are measured according to what is termed *business time*, with $t \in \{1, 2, \ldots, T\}$. Such time does not continue during the lunch period. It will only commence after the lunch period is over. Therefore, the zero increment in time during the lunch period implies that the periodic components in (4.6) do not change over this period. This in turn implies that conditional return volatility before and after the break is equal. This assumption is somewhat restrictive and may not necessarily reflect the actual volatility process before and after the lunch period. As an alternative, Taylor (2004) suggests that the estimation of the periodic components should utilise a measure of time that is based on the actual timing of events. This is referred to as *calendar time*. In order to reflect this more appropriate time measure, the $s(t)$ and S in (4.8) are replaced with $s^c(t)$ and S^c, respectively, where $s^c(t)$ is the calendar time of the tth observation within the periodic cycle, and S^c is the calendar time of the last observation of the periodic cycle. This approach is, henceforth, referred to as *the jointly estimated augmented FFF version of the PGARCH model.*

In order to overcome the somewhat restricted functional form of the FFF version of the PGARCH model, Taylor (2004) introduces the spline-based PGARCH model. This model makes use of cubic spline functions in the estimation of the conditional return volatility and is therefore able to capture complex periodic volatility dynamics. Specifically, the spline-based PGARCH model allows different cubic spline functions to be estimated between selected points (referred to as *knots*) within the periodic cycle. In this instance, we let k_j denote the *j*th knot, with $k_j = \{k_j \in \mathbf{Z}^+ : 0 \le k_j \le S\}, j \in \{0,1,...,J\}$, and $k_0 = 0$. Therefore, if we apply this to (4.2), for example, then the formulation will be as follows:

$$h_t = \sum_{j=1}^{J}\left(\alpha_{1,j}D_j\left(\frac{s(t)-k_j}{S}\right)+\alpha_{2,j}D_j\left(\frac{s(t)-k_j}{S}\right)^2+\alpha_{3,j}D_j\left(\frac{s(t)-k_j}{S}\right)^3\right)+\sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2+\sum_{j=1}^{p}\beta_j h_{t-j} \tag{4.7}$$

Where k_j is the knot position within the periodic cycle, j is the number of knots, and D_j is the coefficient of the spline variable. D_j equals unity if $s(t) \ge k_j$, and zero otherwise. This modelling approach is henceforth referred to as *the jointly estimated spline version of the PGARCH model*. This approach has a similar embedded restriction to the jointly estimated FFF version of the PGARCH model, when an intraday trading break occurs. The use of calendar time instead of business time may allow the periodic components of conditional return volatility to vary between the opening and closing of the trading day. Again, in order to do this, we need to replace *s(t)* and S in (4.7) with $s^c(t)$ and S^c respectively. This approach is henceforth referred to as *the jointly estimated augmented spline version of the PGARCH model.*

4.2.3 Two-step Filtration Periodic GARCH model

An alternative approach to the jointly estimated approaches described above is to apply the two-step filtration approach employed by Andersen and Bollerslev (1997, 1998a) and Martens, Chang, and Taylor (2002). This approach differs in the sense that there is a clear separation in the process of modelling the volatility and estimating the periodicity components. The first step in this approach is to estimate the periodic components of the data used to give the *fitted* periodic components. The next step is to remove these fitted periodic components from the asset returns. This is done by dividing the returns by the fitted periodic components. The final step involves modelling these filtered returns using one's preferred volatility model(s).

The simple intraday means of squared returns are often used to estimate the periodic components (see Martens, Chang and Taylor, 2002, for example). In this study, we estimate the fitted components based on an ordinary least squares (OLS) regression of squared returns on several hourly dummy variables especially created to capture the intraday periodicities present in the dataset. Once the intraday returns data are filtered, we model the volatility using the standard GARCH specifications in (4.2.1.1), (4.2.1.2) and (4.2.1.3). We refer to this approach as *the two-step dummy version of the PGARCH model.* An alternative periodic component that we use to filter the intraday returns is the fitted periodic components from an ordinary least squares regression of squared returns on the FFF-based variables. We call this procedure *the two-step FFF version of the PGARCH model.* The filtered dataset is then modelled again using the various standard GARCH specifications described earlier. Similarly, applying FFF-based variables estimated using the calendar time, we have *the two-step augmented FFF version of the*

PGARCH model. Next we use the spline variables in the regression to obtain the fitted periodic components. Again we divide the returns by these fitted periodic components to remove the periodicity from the return series. We then estimate the conditional return volatility using the five GARCH models explained earlier. This approach is henceforth known as *the two-step spline version of the PGARCH model*. In order to account for the closure during the lunch period, we use periodic components that are measured using the calendar time and apply the two-step filtration method. This approach is therefore known as *the two-step augmented spline version of the PGARCH model.*

4.3 Data, Tests and Results

This section proceeds by describing the data and the methodology used to investigate periodicities in conditional return volatility. It also describes in detail how the thirteen modelling approaches are constructed and used in the volatility estimation. This section concludes by reporting the results of each approach in terms of model fit.

4.3.1 Data

The minute frequency KLCI return data obtained for this study span the period commencing on 29 January 2001 and ending on 29 December 2002 and were obtained from the KLSE. The market is open from Monday through Friday. The morning trading session runs from 9.00am until 12.30 pm, and then closes for a two-hour lunch break. The afternoon trading session then commences at 2.31pm and runs until 5.00pm. The data set gives 146,160 observations, which cover 406 trading days. However, following the recommendations of Andersen and Bollerslev (1998a), we convert the data into five-minute frequency returns. This frequency is deemed to be low enough to avoid stale data, and high enough to avoid loss of information. We compute the 5-minute return intervals as the first logarithmic difference of the index prices measured in percentages terms. Specifically, the 5-minute returns of the KLCI are computed as follows:

$$R_t = 100*\log(p_t / p_{t-1}) \tag{4.8}$$

where R_t is the 5-minute return and p_t is the level of a price index at time t. We obtained 30,044 5-minute return observations and we use these observations in the estimation of all the volatility models described above.

The properties of the KLCI returns for the period under study are presented in Table 4.1. The mean return is negative, indicating that the market is still bearish after experiencing the impact of the Asian financial crisis in 1997-1998. The figures for the maximum, the minimum and the standard deviation of returns over the period are also high, indicating a volatile and unsettled market. The return series skewness coefficient is fairly positive, implying that the distribution of returns is not symmetric but skewed to the right. The coefficient of the sample kurtosis is very high (more than the normal value of 3), indicating that the distribution is highly leptokurtic. The Jarque-Bera statistic for the normality test is highly significant at the 1% level of confidence, suggesting that the null hypothesis of normality can be rejected.

4.3.2 Periodicity Tests

The first stage of the analysis involves an examination of intraday volatility. Figure 4.1 shows the plot of the KLCI returns for the sample period of 406 trading days. The plot clearly demonstrates the volatility clustering effect, which is common in many financial asset returns. Volatility appears to occur in bursts and the plot also shows that the returns are more volatile in the early part of the sample than in the latter part. This is not surprising, considering that the Malaysian economy was still at the recovery stage following the Asian financial crisis in 1997-1998. The situation appears to be improving towards the end of 2002.

Next, we compute the mean intraday value of absolute returns using the entire sample of data. The plot of the mean absolute returns is presented in Figure 4.2. The plot indicates that volatility is high during the first ten minutes and even higher during the last five minutes of the trading period. The plot exhibits the presence of U-shaped patterns for both the morning and afternoon trading sessions. It is interesting to note that there is a surge of volatility five minutes prior to the start of the lunch break. After the lunch break, the volatility appears to drop steadily until it picks up again in the middle of the afternoon session, then drops again before rising to the end of the session. The highest volatility of the trading day occurs a few minutes before the close of trade. A plot of the intraday standard deviation of returns during the day is presented in Figure 4.3. The plot appears to confirm the double U-shaped pattern observed for the morning and afternoon trading sessions. These results are consistent with those found by Andersen, Bollerslev and Cai (2000) using Japanese data, Bildik (2000) using Turkish data, and Taylor (2004) using UK data.

The second stage of the analysis is to design appropriate periodic dummy variables to be incorporated into the conditional equations of the volatility models. We create twelve half-hourly dummies (7 dummies for the morning session and 5 dummies for the afternoon session), which are incorporated into the volatility equations in (4.2), (4.3) and (4.4) above. The dummies are constructed along the following time intervals:

D_1 = 9.01am to 9.30am	D_8 = 2.31pm to 3.00pm
D_2 = 9.31am to 10.00am	D_9 = 3.01pm to 3.30pm
D_3 = 10.01am to 10.30am	D_{10} = 3.31pm to 4.00pm
D_4 = 10.31am to 11.00am	D_{11} = 4.01pm to 4.30pm
D_5 = 11.01am to 11.30am	D_{12} = 4.31pm to 5.00pm
D_6 = 11.31am to 12.00 noon	
D_7 = 12.01pm to 12.30pm	

These dummies are selected and designed to capture the periodicities over the two trading sessions. The lag structure p and q of (1,1) are used in the PGARCH models. This result in 3 competing PGARCH based models. In order to detect the presence of intraday periodicity, we apply Wald tests designed to test for periodicities for each of the models. This is done by restricting the coefficients of the dummy variables to equal zero. The results of the Wald tests are presented in Table 4.2. All results obtained are statistically significant at the 5% level of confidence, indicating the existence of strong intraday periodicities in the KLCI returns. The results suggest the need to consider the impact of periodicity on the dynamic return features when modelling intraday volatility.

4.3.3 Model Estimation

After finding evidence of intraday periodicities in return volatility, specific models of conditional volatility are now estimated. Thirteen different modelling approaches are employed in this study. Each approach, in some form, utilizes the GARCH class of models described in Section 4.2. Specifically, the non-periodic standard conditional volatility models that are used to estimate the conditional variance are the GARCH, the TGARCH and the EGARCH models. All the GARCH-based models are estimated based on the lag structure of (1,1) and by maximizing the quasi-maximum likelihood function, with Bollerslev-Wooldridge robust quasi-maximum likelihood (QML) covariance/standard errors. In addition, all estimation of parameters is carried out using the EViews Version 3.1 software package. We determine the best approach by comparing the model fit produced by the best GARCH-based model in each category.

A brief description of the thirteen different approaches is now given. The first approach is to estimate the KLCI returns data using non-periodic conditional volatility models, i.e., with no periodic components incorporated in the conditional variance equation. This approach is known in this study as the non-periodic GARCH model and this is referred to as approach T1. The next twelve approaches have periodic components incorporated into the conditional variance equation to account for the periodicity in the volatility process. Specifically, the second approach employed half-hourly dummy variables that are equally spaced throughout the trading day. This approach is referred to as the jointly estimated full dummy version of the PGARCH model and is denoted as approach T2. The third approach is the two-step full dummy version of the PGARCH model and is referred to as approach T3. The fourth approach employed 4 quarter-hourly dummy variables which are positioned at the following time intervals:

D_1 = 9.01am to 9.15am $\quad$ D_3 = 2.31pm to 2.45pm

D_2 = 12.15pm to 12.30pm $\quad$ D_4 = 4.45pm to 3.30pm

This approach is therefore referred to as the jointly estimated partial dummy version of the PGARCH model and is denoted as approach T4. The fifth approach is the two-step partial dummy version of the PGARCH model and is referred to as approach T5. The sixth approach is the jointly estimated FFF version of the PGARCH model and is referred to as approach T6. The seventh approach is the two-step FFF version of the PGARCH model and is referred to as approach T7. The eighth approach is the jointly estimated augmented FFF version of the PGARCH model and this is referred to as approach T8. The ninth approach is the two-step augmented FFF version of the PGARCH model and this is referred to as approach T9. The tenth approach is the jointly estimated spline version of the PGARCH model and this is referred to as approach T10. The eleventh approach is the two-step spline version of the PGARCH model and this is referred to as approach T11. The twelfth approach is the jointly estimated augmented spline version of the PGARCH model and this is denoted as approach T12. The final approach is to estimate the conditional return volatility using the two-step augmented spline version of the PGARCH model and this is referred to as approach T13. Please refer to Section 4.2 for detailed descriptions of the basis of each approach.

The first step in T1 is to estimate the three GARCH-based models described above. The estimated parameters obtained for each of the three GARCH-based models are then compared to select the best model fit. Model fit is measured in three ways: the log likelihood (LL), the Akaike Information criterion (AIC) and the Schawrz Information criterion (SIC). However, given the penal nature of the latter two measures, the best model fits are determined by the model that produces the minimum values of AIC and SIC. The estimated parameters for each class, together with their associated Bollerslev-Wooldridge heteroskedastic-consistent standard errors, and measures of model fit, are given in Table 4.3.

The results for the non-periodic GARCH models indicate that the conditional return volatility appears to follow a stationary process $(\hat{\alpha}+\hat{\beta}<1)$, but exhibits a significant degree of time dependency, as indicated by the significant coefficients on past conditional return volatility and past squared errors. The best overall model is the EGARCH model with a LL value of 31981.16, an AIC value of -2.1287 and a SIC value of -2.1273.

The T2 approach follows the same systematic process as above. Each class of GARCH models is now estimated with the twelve half-hourly time-interval dummy variables. These dummy variables are included in the specification of the GARCH class of models in order to account for the intraday volatility process. The estimated parameters are presented in Table 4.4. The results indicate that the T2 approach appears to be competent in capturing the periodicities in the intraday data. In fact, the overall results demonstrate that the T2 approach produces superior results across all classes of GARCH models when compared to the models estimated in T1. This is evinced by the values of the log likelihood functions and the information criteria, which provide a much greater degree of fit than obtained previously. This suggests that the inclusion of periodic components into the variance equations, as demonstrated in the PGARCH structure, does offer a superior description of the volatility dynamics than the non-periodic GARCH models. The results obtained are consistent with the results of previous studies discussed earlier. The most appropriate model is again given by the EGARCH model, with values of LL of 33788.20, AIC of -2.2483, and SIC of -2.2438. One may argue that restricting the periodicity to the intercept term in the conditional volatility equation in this approach seems restrictive. In order to assess whether allowing the intercept to vary is correct, we consider the term α (α is more likely to vary more than the term β in intraday analysis in equation 4.2 above) to also vary over the trading day. In this respect, we multiply each of the twelve dummy variables with the squared error term ε from the previous period and together with the twelve dummy variables, we estimate the return volatility using the GARCH model. The result is shown in Table 4.16. It could be observed that in every respect, this approach is comparable to the T2 approach in terms of model fit.

The T3 approach used in this study is based on the periodicity pattern estimated using intraday squared returns. This involves an OLS regression of squared returns on the twelve dummy variables described above to obtain the fitted periodic components, which are subsequently used to filter the returns. The filtered returns are then used in the parameter estimation for each class of the standard GARCH models employed in the study. The log likelihood (and information criteria) is then adjusted for each case by multiplying the mean adjusted returns by the value of the fitted periodic components and the estimated conditional variances by the squared value of the fitted periodic components. The adjusted log likelihood (and information criteria) for each specification of the GARCH models is then compared to determine the best model fit. The parameter estimates are presented in Table 4.5. The GARCH model gives the best fit, with a LL value of 33308.53, an AIC value of -2.2171, and a SIC value of -2.2159.

The T4 approach is almost similar to the T2 approach. The only difference as mentioned earlier is in the form of two quarter-hourly dummy variables which are positioned at the opening and closing of the trading period respectively and another two which are placed just before and after the lunch time period. The results are presented in Table 4.6. The EGARCH model provides the best model fit with values of LL of

33581.42, AIC of -2.2350 and SIC of -2.2325. Similarly, the mechanics of the T5 approach mirror the mechanics of the T3 approach. The results are reported in Table 4.7. This time around the GARCH model produces the best model fit with values of LL of 33392.07, AIC of -2.2226 and SIC of -2.2215.

The next approach, T6, attempts to model the periodicities in the data with the FFF-based variables instead of the half-hour and quarter-hour dummy variables. This is carried out within the PGARCH structure described in section 4.2.2. This approach uses the business time measurement.[4] We then attempt to find the optimum tuning parameter Q to determine the order of the Fourier expansion. Using a grid search over the space, $q = \{1,\ldots,5\}$, the optimal fit is achieved when the number of FFF variables ($\sin\theta_t$ and $\cos\theta_t$) equals 2, i.e. $Q = 2$. The detailed results are presented in Table 4.8.[5] The results again indicate that periodicity is significant in the data. The best model fit is given by the EGARCH model with a LL value of 33012.92, an AIC value of -2.1971 and a SIC value of -2.1946. The seventh approach, T7, is the two-step FFF version of the PGARCH model. The fitted periodic components are estimated with $Q = 2$, the same as that used in T4. An OLS regression using the four FFF variables is then performed to generate the fitted periodic components, which are subsequently used to produce the filtered returns. The adjusted returns are then used in the parameter estimation for each class of GARCH

---

[4] To clarify the definitions of time, assume that we are using five-minute frequency returns over the trading day, and that trading starts at 9:00 and finishes at 17:00, with a two-hour break in trading between 12:30 and 14:30. This means that the business and calendar times of the last observations of the periodic cycle, S and S^c, will be 72 (= 6 x 12) and 288 (= 24 x 12), respectively. Therefore, at the opening at 9:00, $s(t)/S$ equals 1/72 and $s^c(t)/S^c$ equals 1/288, and at the lunch-time close in trading, $s(t)$ equals 42/72 and $s^c(t)/S^c$ equals 42/288. However, at the opening of trading at 14:30, $s(t)/S$ still equals 42/72, but $s^c(t)/S^c$ has increased to 66/288. It is this difference in the $s^c(t)/S^c$ values at the close of morning trading and the opening of afternoon trading that allows conditional return volatility to be different at these points in time. Similarly, at the close of the trading at 17:00, $s(t)/S$ equals unity while $s^c(t)/S^c$ equals 96/288. Therefore, only the use of the latter ratio will enable the periodic components to differ, and hence, will allow conditional return volatility to differ over these points in time

[5] Results pertaining to other values of Q are available upon request.

model employed in the study. Similarly, the log likelihood is then adjusted for each case by multiplying the mean adjusted returns by the fitted periodic components' value and the estimated conditional variances by the square of the fitted periodic components' value. The results are presented in Table 4.9. The TGARCH model provides the best model fit, with values of LL of 32670.37, AIC of -2.1745 and SIC of -2.1731.

The next approach, T8, is very similar to the technique applied for approach T6. The only difference now is that instead of using business time to model the periodicities in the return volatility, we use the calendar time measurement with $S^c = 288$ and $s^c = 1,\ldots,42, 66,\ldots,96,\ldots,288$ (s^c is the 5-minute return interval). The trading break is indicated by the gap in the 43rd time interval and 66th time interval respectively. For comparison purposes, we apply the same tuning parameter $Q = 2$, which is used in approaches T6 and T7, in order to determine the order of the Fourier expansion. The results are reported in Table 4.10. The best model fit is again given by the EGARCH model, with a LL value of 32652.07, an AIC value of -2.1731 and a SIC value of -2.1706. The technique applied in approach T9 is similar to the one used in approach T7. The difference lies in the use of the calendar time measurement, which has been described in approach T8 above, instead of the business time measurement. Again, the fitted periodic components are estimated with $Q = 2$, the same as that used in approaches T6, T7 and T8 above. The results are presented in Table 4.11. The TGARCH model provides the best model fit with values of LL of 32564.72, AIC of -2.1675 and SIC of -2.1661.

The next approach, T10, attempts to model the periodicities in the data with the spline-based variables. This approach uses the business time measurement with $S = 72$ and $s = 1,\ldots,72$ (s is the 5-minute return interval). In order to estimate the periodic components, we need to select the appropriate number and position of knots to obtain the optimal AIC statistics. Based on the number of observed 5-minute time intervals, we assume that four (approximately) equally spaced intraday knots occur at the following positions: at $k_0 = 0$, $k_1 = 19$, $k_2 = 37$ and $k_3 = 56$ respectively. We then incorporate these knots into the estimation of the conditional variance as formulated in equation 4.7 for each of the three GARCH models. The results are reported in Table 4.12. The results indicate that periodicity is significant in the data. The best model fit is given by the EGARCH model with a LL value of 33698.65, an AIC value of -2.2422 and a SIC value of -2.2375. The eleventh approach, T11, is the spline version of the two-step filtration technique. The fitted periodic components are estimated with the four knots identified in approach T6. An OLS regression is then performed to generate the fitted periodic components, which are subsequently used to produce the filtered returns. The adjusted returns are then used in the parameter estimation for each of the five standard GARCH models. The log likelihood is then adjusted for each case by multiplying the mean adjusted returns by the fitted periodic components' value and the estimated conditional variances by the square of the fitted periodic components' value. The results are presented in Table 4.13. The TGARCH model provides the best model fit, with values of LL of 33551.49, AIC of -2.2332, and SIC of -2.2318.

The twelfth approach, T12, and the thirteenth approach, T13, are similar in almost all aspects to the approaches T10 and T11 respectively. However, the T12 and T13 approaches make use of calendar time. Unlike T10 and T11, for both T12 and T13, the positions of the knots are assumed to be at $k_0 = 0$, $k_1 = 24$, $k_2 = 48$ and $k_3 = 72$

respectively.[6] Similar estimation techniques to those described in approaches T10 and T11 are then applied. The results for the T12 are reported in Table 4.14. The best model fit is again given by the EGARCH model with a LL value of 33736.84, an AIC value of -2.2447 and an SIC value of -2.2399. The results for the T13 approach are presented in Table 4.15. The TGARCH model provides the best model fit, with values of LL of 33536.20, AIC of -2.2321 and SIC of -2.2308.

The relative performances of all the modelling approaches described above are presented in Table 4.164. The GARCH model with the best model fits for each approach is reported, together with the corresponding LL, AIC and SIC statistics. The performance of the approaches is then ranked based on the AIC (for in-sample evaluation) and the SIC (for out-of-sample evaluation) statistics. It is clear that based on the AIC and SIC rankings, the best performing approach appears to be the T2 approach. This is followed by T12, T10, T4, T11, T13, T5, T3, T6, T7, T8, T9 and finally T1. It is clear from Table 4.16 that any modelling approach that accounts for periodicity produces superior results to the non-periodic approach. This suggests that the PGARCH structure provides a better explanation and superior information regarding the periodicity effects in intraday conditional volatility. The overall results suggest that the best approach to estimate the conditional return volatility for the KLCI returns is to apply approach T2, that is, to jointly estimate the half-hourly dummy variables in the conditional variance equation. However, if we are to group the periodic approaches according to the types of variable

[6] Different positions of the knots are selected because the calendar time measurement is used, i.e., S^c = 288, instead of the business time measurement, where S = 72. The usage of the calendar time measurement is useful as it allows the modelling of the discontinuity in conditional return volatility during trading breaks. Time, therefore, does increase, and periodic components do change during the break, implying that conditional return volatility before and after the break will not be the same. The different positions of the knots for T10 and T11 are therefore different, due to the longer time period measurement.

used in the estimation, there is a strong indication that approaches that incorporate spline-based variables generally perform better than the approaches that employ half-hourly dummy variables and the FFF-based variables, respectively. This is true for both AIC and SIC rankings, if we exclude the performance of the T2 approach.

It is also difficult to establish whether the jointly estimated technique is superior to the two-step filtration technique. For example, for the dummy version of the PGARCH models, it is apparent that the jointly estimated technique is superior to the two-step filtration technique. For the FFF version of the PGARCH models, similar finding is observed. For modelling based on business and calendar time measurement, the jointly estimated based technique is superior to the two-step filtration technique. It is also observed that modelling based on business time produces superior model fit than modelling based on calendar time (the model fit of approaches T6 and T7 are better than the model fit of approaches T8 and T9). Approach T6 provides the best overall result for the FFF version of the PGARCH models. For the spline version of the PGARCH models, the position is much clearer. The jointly estimated technique gives superior model fit over the two-step filtration technique when both the business time and calendar time measurements are applied. The best results for the spline-based variables is produced by T10, which is a jointly estimated technique using the calendar time measurement. There is, however, no clear evidence suggesting the superiority of approaches that utilize the calendar time measurement over the approaches that are based on the business time measurement.

Another important finding from the observation based on Tables 4.3 to 4.15 is that the EGARCH model specification produced consistently superior results to other

GARCH specifications used in all thirteen approaches. This suggests that modelling intraday conditional volatility, at least for the KLSE, using the EGARCH could provide a better explanation for the asymmetric relationship between returns and volatility changes.

The performance and the ability of each approach (T1 to T13) to capture features of the intraday volatility periodicity can be examined further by inspection of Figures 4.4 to 4.10. A clear periodicity is apparent from the plots. The return volatility is found to be high during the opening of trading and the time just prior to the lunch hour. Volatility is also high at the opening of trading after the lunch hour. The highest volatility occurs during the last five minutes of trading. From Figure 4.10, it is clear that the T2 approach produces the best volatility fit when compared against the other ten competing approaches. It is also clear from Figures 4.4 to 4.9 that in all cases, the PGARCH models produced superior volatility fit to the non-periodic GARCH models. This confirms the findings above.

4.4 Conclusion

This study provides a detailed investigation into intraday volatility dynamics in the Malaysian stock market. The data used are based on 5-minute frequency returns of the KLCI series. Two types of test are conducted. The first focuses on the intraday volatility pattern or periodicity of returns across the trading day, taking into account the closure for the lunch break. Consistent with previous studies, we find that the intraday volatility is dominated by two separate U-shaped patterns: one for the morning trading session and another in the afternoon trading session. The heightened volatility around the opening and closing of the two separate trading sessions on the KLSE is broadly consistent with the predictions from theoretical market microstructure models based on the strategic interaction of asymmetrically informed agents suggested by Admati and Pfleiderer (1988) and Foster and Viswanathan (1990).

The second test is intended to provide insights into the methods of modelling the intraday volatility periodicity on the KLSE. The results generally indicate that modelling approaches that incorporate periodic components in estimating the conditional variance provide a greater degree of model fit and better performance compared to the performance of the non-periodic conditional volatility models. We compared the performance of the two-step filtration technique of Andersen and Bollerslev (1997, 1998a) with the jointly estimated technique, both within the PGARCH structure suggested by Bollerslev and Ghysels (1996), using four types of variables, namely half-hourly and quarter-hourly dummy variables, FFF-based variables and spline-based variables. These variables are designed to capture the periodicity effect in the returns data. Consistent with the findings of Martens, Chang and Taylor (2002), we find some evidence that the jointly estimated technique does provide superior performance over the two-step filtration technique. This is the case when we find that the jointly estimated full dummy version of the PGARCH model approach produces the best performance among the thirteen approaches evaluated in this study. The jointly estimated technique is clearly dominant when half-hour and quarter-hour dummy variables and spline-based variables are used in the estimation of the conditional volatility equations. We also find that the two-step filtration approaches incorporating the spline-based variables and FFF-based variables offer an encouraging and less (computationally) expensive alternative to the jointly estimated modelling approaches. For example, both the two-step spline version and the two-step augmented spline version of the PGARCH models produced comparable performance to the more expensive approach of the jointly estimated spline version and the jointly estimated augmented spline version of the PGARCH models evaluated in the study.

The overall results of this chapter support the case that the practical estimation and extraction of the intraday periodic component of return volatility is both feasible and indispensable for a meaningful intraday dynamic analysis. We evaluate whether modelling the intraday conditional volatility using the calendar time measurement, as suggested by Taylor (2004), offers any significant advantage over the business time measurement. The results are somewhat mixed, but a very encouraging result is shown by the jointly estimated augmented spline version of the PGARCH model approach, which uses the calendar time measurement. The gap in performance is very small when this approach is compared using AIC and SIC rankings in relation to the best approach, which is the jointly estimated full dummy version of the PGARCH model.

Finally, the results show that at least for the GARCH-based models, there is a motivation for using the EGARCH model to accommodate the asymmetry in the relationship between returns and volatility changes. Results for the jointly estimated based approaches indicate that the EGARCH model consistently produce superior model fit compared to the other GARCH-based models used in the study.

The findings from this chapter could provide a clue to the expected forecasting performance of the thirteen modelling approaches, which will be discussed in detail in Chapter 5. Findings from previous studies suggest that the success of a volatility modelling approach lies in its out-of-sample forecasting power. Therefore, it would be interesting to see whether the jointly estimated full dummy version of the PGARCH model approach could continue to produce superior performance among the thirteen approaches. The same could be expected for the modelling approaches that utilized spline-based variables, which have also shown strong in-sample performances. We would also like to see whether the two-step filtration based modelling approaches (which are computationally less expensive) could produce superior forecasting performance compared with the jointly estimated based modelling approaches. In addition, it would be interesting to assess the accuracy of the forecasts from the thirteen modelling approaches in mapping the *ex post* realized volatility, which we suspect will exhibit the double U-shaped intraday periodicity pattern.

Table 4.1: Summary Statistics for the KLCI Returns

This table reports the various statistics for the KLCI returns. The period under examination is from 29 January 2001 to 29 December 2002

	KLCI Returns
Mean	-0.0004
Standard Deviation	0.1131
Skewness	0.8621
Kurtosis	183.1164
Maximum	3.5741
Minimum	-4.37351
Jarque-Bera	40614251
	(0.0000)[1]

[1] The number in brackets is the *p*-value for the corresponding Jarque-Bera statistic.

Table 4.2: Testing for Periodicity

This table contains the F-statistics under the Wald tests associated with the joint estimate using the PGARCH periodicity test for intraday periodicity $(F(\tilde{D}_d = 0))$. The dummies are constructed along the following time intervals:

D_1 = 9.01am to 9.30am	D_8 = 2.31pm to 3.00pm
D_2 = 9.31 am to 10.00am	D_9 = 3.01pm to 3.30pm
D_3 = 10.01am to 10.30am	D_{10} = 3.31pm to 4.00pm
D_4 = 10.31am to 11.00am	D_{11} = 4.01pm to 4.30pm
D_5 = 11.01am to 11.30am	D_{12} = 4.31pm to 5.00pm
D_6 = 11.31am to 12.00 noon	
D_7 = 12.01pm to 12.30pm	

These tests are based on 5-minute returns frequency of the KLCI returns (R_t). These tests are conducted for the sample period commencing on 29 January 2001 and ending on 29 December 2002. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

Model	Periodicity Test $(F(\tilde{D}_d = 0))$
GARCH	6.3992**
TGARCH	6.4026**
EGARCH	4.7305**

Table 4.3: Comparison of the Non-Periodic GARCH models – KLCI

This table describes the parameter estimates of the Non-Periodic GARCH models described and denoted in the text as approach T1. The numbers in parentheses are Bollerslev-Wooldridge robust standard errors. The log likelihood (LL), the Akaike Information Criterion (AIC), and the Schwarz Information Criterion (SIC) are also given. The significance of these estimates is denoted by ** (1% significance) and * (5% significance).

	GARCH	TGARCH	EGARCH
Mean Equation			
C	-0.0004**	-0.0006**	-0.0008
	(0.0005)	(0.0004)	(0.0005)
Volatility Equation			
α_1	0.1527*	0.1467**	
	(0.0110)	(0.0120)	
β_1	0.8042**	0.8031**	
	(0.0104)	(0.0104)	
(RESID<0)* α_1		0.0137	
		(0.01431)	
|RES|/SQR[GARCH](1)			0.2492**
			(0.0180)
RES/SQR[GARCH](1)			-0.0020
			(0.0083)
EGARCH(1)			0.9487**
			(0.0058)
C	0.0005**	0.0005**	-0.0008**
	(5.07E-05)	(5.08E-05)	(0.0005)
Model Fit			
LL	31952.65	31954.42	31981.56
AIC	-2.1269	-2.1269	-2.1287
SIC	-2.1258	-2.1255	-2.1273

Table 4.4: Comparison of the Jointly Estimated Full Dummy version of the PGARCH Model – KLCI

This table describes the parameter estimates of the PGARCH models described and denoted in the text as approach T2. The numbers in parentheses are Bollerslev-Wooldridge robust standard errors. B_1 to B_{11} are the coefficients of the dummy variables described in Table 4.2 above. The log likelihood (LL), the Akaike Information Criterion (AIC), and the Schwarz Information Criterion (SIC) are also given. The significance of these estimates is denoted by ** (1% significance) and * (5% significance).

	GARCH	TGARCH	EGARCH
Mean Equation			
C	-0.0013**	-0.0013**	-0.0013**
	(0.0004)	(0.0004)	(0.0004)
Volatility Equation			
α_1	0.0432**	0.0428**	
	(0.0036)	(0.0042)	
β_1	0.9370**	0.9374**	
	(0.0048)	(0.0048)	
(RESID<0)* α_1		0.0003	
		(0.0038)	
|RES|/SQR[GARCH](1)			0.0673**
			(0.0073)
RES/SQR[GARCH](1)			-0.0002
			(0.0032)
EGARCH(1)			0.9956**
			(0.0009)
B_1	-0.0039**	-0.0039**	-0.5152**
	(0.0009)	(0.0009)	(0.0311)
B_2	-0.0027**	-0.0026**	-0.3696**
	(0.0005)	(0.0004)	(0.0227)
B_3	-0.0025**	-0.0025**	-0.3044**
	(0.0005)	(0.0005)	(0.0253)
B_4	-0.0027**	-0.0027**	-0.3789**
	(0.0005)	(0.0005)	(0.0252)
B_5	-0.0025**	-0.0025**	-0.2872**
	(0.0005)	(0.0005)	(0.0242)
B_6	-0.0027**	-0.0027**	-0.3778**
	(0.0005)	(0.0005)	(0.0254)
B_7	-0.0021**	-0.0020**	-0.1753**
	(0.0005)	(0.0005)	(0.0378)
B_8	-0.0027**	-0.0027**	-0.3730**
	(0.0005)	(0.0005)	(0.0332)
B_9	-0.0026**	-0.0025**	-0.3361**
	(0.0005)	(0.0005)	(0.0249)
B_{10}	-0.0024**	-0.0024**	-0.2902**
	(0.0005)	(0.0005)	(0.0232)
B_{11}	-0.0026**	-0.0026**	-0.3512**
	(0.0005)	(0.0005)	(0.0300)
C	0.0026**	0.0026**	0.2467**
	(0.0005)	(0.0005)	(0.0195)
Model Fit			
LL	33369.82	33369.89	33788.20
AIC	-2.2205	-2.2204	-2.2483
SIC	-2.2163	-2.2160	-2.2438

Table 4.5: Comparison of the Two-step Full Dummy version of the PGARCH Models – KLCI

This table describes the parameter estimates of the PGARCH models described and denoted in the text as approach T3. The numbers in parentheses are Bollerslev-Wooldridge robust standard errors. The log likelihood (LL), the Akaike Information Criterion (AIC), and the Schwarz Information Criterion (SIC) are also given. The log likelihood is based on the data used to estimate each GARCH model specification. Since R^2 is used to estimate the periodicity pattern, the adjusted log likelihood is reported for each case. The adjusted log likelihood is obtained by multiplying the residuals by the periodicity pattern as well as multiplying the estimated conditional variances by the square of the periodicity term. The significance of these estimates is denoted by ** (1% significance) and * (5% significance).

	GARCH	TGARCH	EGARCH
Mean Equation			
C	-0.0075	-0.0077	-0.0069
	(0.0051)	(0.0049)	(0.0052)
Volatility Equation			
α_1	0.0226**	0.0224**	
	(0.0022)	(0.0028)	
β_1	0.9741**	0.9739**	
	(0.0022)	(0.0022)	
(RESID<0)* α_1		0.0007	
		(0.0031)	
|RES|/SQR[GARCH](1)			0.0620**
			(0.0079)
RES/SQR[GARCH](1)			-0.0009
			(0.0037)
EGARCH(1)			0.9957**
			(0.0009)
C	0.0031**	0.0031**	-0.0445**
	(0.0008)	(0.0008)	(0.0057)
Model Fit			
LL	-39173.43	-39173.34	-39229.18
Adjusted LL	33308.53	33308.62	33252.79
AIC	2.6081	2.6082	2.6119
Adjusted AIC	-2.2171	-2.2169	-2.2133
SIC	2.6092	2.6095	2.6133
Adjusted SIC	-2.2159	-2.2156	-2.2119

Table 4.6: Comparison of the Jointly Estimated Partial Dummy version of the PGARCH Model – KLCI

This table describes the parameter estimates of the PGARCH models described and denoted in the text as approach T4. The numbers in parentheses are Bollerslev-Wooldridge robust standard errors. B_1 to B_4 are the coefficients of the dummy variables described as below respectively:

D_1 = 9.01am to 9.15am
D_2 = 12.15 pm to 12.30 pm
D_3 = 14.31 pm to 14.45 pm
D_4 = 16.45 pm to 17.00 pm

The log likelihood (LL), the Akaike Information Criterion (AIC), and the Schwarz Information Criterion (SIC) are also given. The significance of these estimates is denoted by ** (1% significance) and * (5% significance).

	GARCH	TGARCH	EGARCH
Mean Equation			
C	-0.0011*	-0.0012*	-0.0012**
	(0.0005)	(0.0005)	(0.0004)
Volatility Equation			
α_1	0.1479**	0.1473**	
	(0.0184)	(0.0195)	
β_1	0.5967**	0.5966**	
	(0.0416)	(0.0366)	
(RESID<0)* α_1		0.0466	
		(0.0250)	
|RES|/SQR[GARCH](1)			0.1455**
			(0.0111)
RES/SQR[GARCH](1)			0.0021
			(0.0052)
EGARCH(1)			0.9743**
			(0.0034)
B_1	-0.0005	-0.0005	-0.3149**
	(0.0006)	(0.0007)	(0.0235)
B_2	-0.0001	-0.0001	0.2306**
	(0.0005)	(0.0005)	(0.0518)
B_3	-0.0002	-0.0002	-0.0579
	(0.0004)	(0.0004)	(0.0444)
B_4	-0.0001	-0.00003	0.5595**
	(0.0010)	(0.0010)	(0.0242)
C	0.0090**	0.0090**	-0.2583**
	(0.0009)	(0.0008)	(0.0221)
Model Fit			
LL	23050.44	22949.12	33581.42
AIC	-1.5340	-1.5272	-2.2350
SIC	-1.5318	-1.5247	-2.2325

Table 4.7: Comparison of the Two-step Partial Dummy version of the PGARCH Models – KLCI

This table describes the parameter estimates of the PGARCH models described and denoted in the text as approach T5. The numbers in parentheses are Bollerslev-Wooldridge robust standard errors. The log likelihood (LL), the Akaike Information Criterion (AIC), and the Schwarz Information Criterion (SIC) are also given. The log likelihood is based on the data used to estimate each GARCH model specification. Since R^2 is used to estimate the periodicity pattern, the adjusted log likelihood is reported for each case. The adjusted log likelihood is obtained by multiplying the residuals by the periodicity pattern as well as multiplying the estimated conditional variances by the square of the periodicity term. The significance of these estimates is denoted by ** (1% significance) and * (5% significance).

	GARCH	TGARCH	EGARCH
Mean Equation			
C	-0.0094	-0.0092	-0.0083
	(0.0050)	(0.0049)	(0.0051)
Volatility Equation			
α_1	0.0262**	0.0265**	
	(0.0024)	(0.0031)	
β_1	0.9699**	0.9699**	
	(0.0024)	(0.0024)	
(RESID<0)* α_1		-0.0006	
		(0.0033)	
|RES|/SQR[GARCH](1)			0.0711**
			(0.0087)
RES/SQR[GARCH](1)			0.0007
			(0.0038)
EGARCH(1)			0.9951**
			(0.0011)
C	0.0036**	0.0036**	-0.0512**
	(0.0009)	(0.0009)	(0.0063)
Model Fit			
LL	-38987.65	-38987.57	-39051.60
Adjusted LL	33392.07	33392.15	33328.12
AIC	2.5957	2.5958	2.6000
Adjusted AIC	-2.2226	-2.2225	-2.2184
SIC	2.5968	2.5972	2.6014
Adjusted SIC	-2.2215	-2.2212	-2.2167

Table 4.8: Comparison of the Jointly Estimated FFF version of the PGARCH models – KLCI

This table describes the parameter estimates of the PGARCH models described and denoted in the text as approach T6. The numbers in parentheses are Bollerslev-Wooldridge robust standard errors. The log likelihood (LL), the Akaike Information Criterion (AIC), and the Schwarz Information Criterion (SIC) are also given. The significance of these estimates is denoted by ** (1% significance) and * (5% significance).

	GARCH	TGARCH	EGARCH
Mean Equation			
C	-0.0002	-0.0012**	-0.0019**
	(0.0004)	(0.0004)	(0.0004)
Volatility Equation			
α_1	0.0783**	0.0771**	
	(0.0058)	(0.0065)	
β_1	0.8826**	0.8848**	
	(0.0069)	(0.0066)	
(RESID<0)* α_1		0.0039	
		(0.0073)	
|RES|/SQR[GARCH](1)			0.1203**
			(0.0109)
RES/SQR[GARCH](1)			0.0017
			(0.0051)
EGARCH(1)			0.9877**
			(0.0021)
Var 1	0.0002**		0.0068**
	(2.67E-05)		(0.0026)
Var 2	-0.0002**		-0.0356**
	(1.40E-05)		(0.0023)
Var 3	0.0003**		0.0439**
	(3.58E-03)		(0.0052)
Var 4	-0.0001**		-0.0469**
	(2.17E-05)		(0.0038)
C	0.0003**	0.0003**	-0.1466**
	(2.80E-05)	(2.72E-05)	(0.0145)
Model Fit			
LL	32618.42	32625.42	33012.92
AIC	-2.1709	-2.1713	-2.1971
SIC	-2.1687	-2.1688	-2.1946

Table 4.9: Comparison of the Two-step FFF version of the PGARCH models – KLCI

This table describes the parameter estimates of the PGARCH models described and denoted in the text as approach T7. The numbers in parentheses are Bollerslev-Wooldridge robust standard errors. The log likelihood (LL), the Akaike Information Criterion (AIC), and the Schwarz Information Criterion (SIC) are also given. The log likelihood is based on the data used to estimate each GARCH model specification. Since R^2 is used to estimate the periodicity pattern, the adjusted log likelihood is reported for each case. The adjusted log likelihood is obtained by multiplying the residuals by the periodicity pattern as well as multiplying the estimated conditional variances by the square of the periodicity term. The significance of these estimates is denoted by ** (1% significance) and * (5% significance).

	GARCH	TGARCH	EGARCH
Mean Equation			
C	-0.0085	-0.0104*	-0.0088
	(0.0051)	(0.0049)	(0.0051)
Volatility Equation			
α_1	0.0279**	0.0261**	
	(0.0024)	(0.0029)	
β_1	0.9679**	0.9699**	
	(0.0027)	(0.0025)	
(RESID<0)* α_1		0.0004	
		(0.0035)	
|RES|/SQR[GARCH](1)			0.0731**
			(0.0081)
RES/SQR[GARCH](1)			-0.0006
			(0.0040)
EGARCH(1)			0.9947**
			(0.0011)
C	0.0044**	0.0039**	-0.0517**
	(0.0009)	(0.0008)	(0.0058)
Model Fit			
LL	-40036.18	-40035.05	-40093.75
Adjusted LL	32669.27	32670.37	32611.67
AIC	2.6655	2.6655	2.6694
Adjusted AIC	-2.1745	-2.1745	-2.1706
SIC	2.6666	2.6669	2.6708
Adjusted SIC	-2.1734	-2.1731	-2.1692

Table 4.10: Comparison of the Jointly Estimated Augmented FFF version of the PGARCH models – KLCI

This table describes the parameter estimates of the PGARCH models described and denoted in the text as approach T8. The numbers in parentheses are Bollerslev-Wooldridge robust standard errors. The log likelihood (LL), the Akaike Information Criterion (AIC), and the Schwarz Information Criterion (SIC) are also given. The significance of these estimates is denoted by ** (1% significance) and * (5% significance).

	GARCH	TGARCH	EGARCH
Mean Equation			
C	-0.0024**	-0.0016**	-0.0014**
	(0.0004)	(0.0004)	(0.0004)
Volatility Equation			
α_1	0.0991**	0.0949**	
	(0.0068)	(0.0075)	
β_1	0.8598**	0.8591**	
	(0.0075)	(0.0075)	
(RESID<0)* α_1		0.0079	
		(0.0088)	
|RES|/SQR[GARCH](1)			0.1969**
			(0.0161)
RES/SQR[GARCH](1)			-0.0026
			(0.0072)
EGARCH(1)			0.9477**
			(0.0064)
Var 1	7.52E-05	6.88E-05	0.0036
	(4.15E-05)	(4.11E-05)	(0.0087)
Var 2	-3.51E-05*	-3.37E-05*	-0.0094*
	(1.54E-05)	(1.54E-05)	(0.0045)
Var 3	0.0003**	0.0003**	0.0644**
	(4.66E-05)	(4.60E-05)	(0.0090)
Var 4	-0.0002**	-0.0002**	-0.0517**
	(2.86E-05)	(2.86E-05)	(0.0064)
C	0.0004**	0.0004**	-0.4107**
	(3.85E-05)	(3.83E-05)	(0.0399)
Model Fit			
LL	32571.01	32576.40	32652.07
AIC	-2.1678	-2.1681	-2.1731
SIC	-2.1656	-2.1656	-2.1706

Table 4.11: Comparison of the Two-step Augmented FFF version of the PGARCH models – KLCI

This table describes the parameter estimates of the PGARCH models described and denoted in the text as approach T9. The numbers in parentheses are Bollerslev-Wooldridge robust standard errors. The log likelihood (LL), the Akaike Information Criterion (AIC), and the Schwarz Information Criterion (SIC) are also given. The log likelihood is based on the data used to estimate each GARCH model specification. Since R^2 is used to estimate the periodicity pattern, the adjusted log likelihood is reported for each case. The adjusted log likelihood is obtained by multiplying the residuals by the periodicity pattern as well as multiplying the estimated conditional variances by the square of the periodicity term. The significance of these estimates is denoted by ** (1% significance) and * (5% significance).

	GARCH	TGARCH	EGARCH
Mean Equation			
C	-0.0097	-0.0099*	-0.0086
	(0.0049)	(0.0049)	(0.0049)
Volatility Equation			
α_1	0.0307**	0.0279**	
	(0.0025)	(0.0029)	
β_1	0.9644**	0.9675**	
	(0.0029)	(0.0027)	
(RESID<0)* α_1		0.0007	
		(0.0035)	
|RES|/SQR[GARCH](1)			0.0753**
			(0.0076)
RES/SQR[GARCH](1)			-0.0008
			(0.0037)
EGARCH(1)			0.9944**
			(0.0011)
C	0.0052**	0.0045**	-0.0531**
	(0.0009)	(0.0009)	(0.0054)
Model Fit			
LL	-40090.39	-40088.53	-40131.49
Adjusted LL	32562.86	32564.72	32521.76
AIC	2.6691	2.6691	2.6719
Adjusted AIC	-2.1674	-2.1675	-2.1646
SIC	2.6702	2.6705	2.6733
Adjusted SIC	-2.1663	-2.1661	-2.1632

Table 4.12: Comparison of the Jointly Estimated Spline version of the PGARCH models – KLCI

This table describes the parameter estimates of the PGARCH models described and denoted in the text as approach T10. The numbers in parentheses are Bollerslev-Wooldridge robust standard errors. The log likelihood (LL), the Akaike Information Criterion (AIC), and the Schwarz Information Criterion (SIC) are also given. The significance of these estimates is denoted by ** (1% significance) and * (5% significance).

	GARCH	TGARCH	EGARCH
Mean Equation			
C	-0.0012**	-0.0013**	-0.0012**
	(0.0004)	(0.0004)	(0.0004)
Volatility Equation			
α_1	0.0633**	0.0625**	
	(0.0013)	(0.0016)	
β_1	0.9061**	0.9064**	
	(0.0015)	(0.0016)	
(RESID<0)* α_1		0.0013	
		(0.0021)	
|RES|/SQR[GARCH](1)			0.0705**
			(0.0012)
RES/SQR[GARCH](1)			-0.0003
			(0.0009)
EGARCH(1)			0.9949**
			(0.0002)
K1	0.1404*	0.1481*	22.7240*
	(0.0602)	(0.0602)	(11.3708)
K2	-4.0279	-4.7234	-440.8920
	(6.7163)	(6.7149)	(1254.2250)
K3	-3.9563	13.1251	302.1174
	(211.9122)	(211.8415)	(39996.65)
K4	0.0817	0.0887	-20.1064
	(0.0564)	(0.0564)	(13.3339)
K5	-4.2439	-5.5737	1395.085
	(4.3829)	(4.3813)	(923.8201)
K6	301.0686	325.5863	271.2493
	(288.5744)	(288.6183)	(63386.32)
K7	0.1705**	0.1518**	-38.1993**
	(0.0471)	(0.0473)	(13.2347)
K8	-41.2169**	-41.3318**	-316.5800
	(2.5079)	(2.5118)	(848.8320)
K9	905.9667**	831.7836**	212.9649
	(199.3422)	(200.3341)	(56702.69)
K10	-0.2449**	-0.2502**	-36.9277**
	(0.0575)	(0.0578)	(13.7952)
K11	-52.6003**	-50.0337**	2033.9910*
	(5.0038)	(5.0331)	(1027.4290)
K12	454.1034	406.1318	98.8022
	(296.9193)	(298.6590)	(64401.51)
C	-0.0012**	-0.0012**	-0.3385**
	(0.0001)	(0.0001)	(0.0277)
Model Fit			
LL	33284.87	33281.73	33698.65
AIC	-2.2147	-2.2144	-2.2422
SIC	-2.2103	-2.2097	-2.2375

Table 4.13: Comparison of the Two-Step Spline version of the PGARCH models – KLCI

This table describes the parameter estimates of the PGARCH models described and denoted in the text as approach T11. The numbers in parentheses are Bollerslev-Wooldridge robust standard errors. The log likelihood (LL), the Akaike Information Criterion (AIC), and the Schwarz Information Criterion (SIC) are also given. The log likelihood is based on the data used to estimate each GARCH model specification. Since R^2 is used to estimate the periodicity pattern, the adjusted log likelihood is reported for each case. The adjusted log likelihood is obtained by multiplying the residuals by the periodicity pattern as well as multiplying the estimated conditional variances by the square of the periodicity term. The significance of these estimates is denoted by ** (1% significance) and * (5% significance).

	GARCH	TGARCH	EGARCH
Mean Equation			
C	-0.0113*	-0.0111*	-0.0096
	(0.0050)	(0.0049)	(0.0052)
Volatility Equation			
α_1	0.0247**	0.0249**	
	(0.0022)	(0.0028)	
β_1	0.9717**	0.9718**	
	(0.0022)	(0.0022)	
(RESID<0)* α_1		-0.0006	
		(0.0031)	
|RES|/SQR[GARCH](1)			0.0646**
			(0.0078)
RES/SQR[GARCH](1)			0.0003
			(0.0034)
EGARCH(1)			0.9955**
			(0.0009)
C	0.0035**	0.0034**	-0.0465**
	(0.0008)	(0.0008)	(0.0056)
Model Fit			
LL	-39760.52	-39760.44	-39842.10
Adjusted LL	33551.40	33551.49	33469.83
AIC	2.6471	2.6471	2.6526
Adjusted AIC	-2.2332	-2.2332	-2.2277
SIC	2.6482	2.6485	2.6539
Adjusted SIC	-2.2321	-2.2318	-2.2263

Table 4.14: Comparison of the Jointly Estimated Augmented Spline version of the PGARCH models – KLCI

This table describes the parameter estimates of the PGARCH models described and denoted in the text as approach T12. The numbers in parentheses are Bollerslev-Wooldridge robust standard errors. The log likelihood (LL), the Akaike Information Criterion (AIC), and the Schwarz Information Criterion (SIC) are also given. The significance of these estimates is denoted by ** (1% significance) and * (5% significance).

	GARCH	TGARCH	EGARCH
Mean Equation			
C	-0.0013**	-0.0013**	-0.0015**
	(0.0004)	(0.0004)	(0.0004)
Volatility Equation			
α_1	0.0625**	0.0621**	
	(0.0013)	(0.0015)	
β_1	0.9081**	0.9079**	
	(0.0015)	(0.0015)	
(RESID<0)* α_1		0.0009	
		(0.0021)	
|RES|/SQR[GARCH](1)			0.0703**
			(0.0013)
RES/SQR[GARCH](1)			-0.0003
			(0.0009)
EGARCH(1)			0.9952**
			(0.0002)
K1	0.1536**	0.1534**	26.3727**
	(0.0230)	(0.0232)	(6.1101)
K2	-5.9993**	-5.9852**	-675.7031
	(2.1298)	(2.1446)	(545.3857)
K3	77.9203	77.6183	495.1117
	(55.0443)	(55.3722)	(14050.96)
K4	0.0203	0.0200	-2.0827
	(0.0354)	(0.0355)	(10.0023)
K5	-6.2341*	-6.1712**	1516.424*
	(2.4278)	(2.4284)	(697.8170)
K6	355.8842**	354.0881*	351.0654
	(136.1915)	(136.6791)	(39742.53)
K7	-0.3189	-0.3155	-99.9645
	(0.3606)	(0.3609)	(71.8476)
K8	-55.4306	-55.4558	1215.923
	(33.5288)	(33.5448)	(5920.305)
K9	414.1275	417.9437	149.2934
	(769.2131)	(769.8277)	(146067.2)
K10	0.0706	0.0724	-65.0524**
	(0.1124)	(0.1124)	(22.7732)
K11	-59.4579**	-59.5813	-694.4943
	(21.3565)	(21.3681)	(3775.867)
K12	331.0778	333.1204	117.4853
	(790.9784)	(791.4737)	(147248.4)
C	-0.0013**	-0.0012**	-0.3426**
	(5.74E-05)	(5.77E-05)	(0.0184)
Model Fit			
LL	33309.42	33309.38	33736.84
AIC	-2.2163	-2.2162	-2.2447
SIC	-2.2119	-2.2115	-2.2399

Table 4.15: Comparison of the Two-Step Augmented Spline version of the PGARCH models – KLCI

This table describes the parameter estimates of the PGARCH models described and denoted in the text as approach T13. The numbers in parentheses are Bollerslev-Wooldridge robust standard errors. The log likelihood (LL), the Akaike Information Criterion (AIC), and the Schwarz Information Criterion (SIC) are also given. The log likelihood is based on the data used to estimate each GARCH model specification. Since R^2 is used to estimate the periodicity pattern, the adjusted log likelihood is reported for each case. The adjusted log likelihood is obtained by multiplying the residuals by the periodicity pattern as well as multiplying the estimated conditional variances by the square of the periodicity term. The significance of these estimates is denoted by ** (1% significance) and * (5% significance).

	GARCH	TGARCH	EGARCH
Mean Equation			
C	-0.0133**	-0.0127**	-0.0108*
	(0.0049)	(0.0049)	(0.0051)
Volatility Equation			
α_1	0.0249**	0.0256**	
	(0.0024)	(0.0029)	
β_1	0.9716**	0.9718**	
	(0.0025)	(0.0024)	
(RESID<0)* α_1		-0.0018	
		(0.0031)	
|RES|/SQR[GARCH](1)			0.0638**
			(0.0084)
RES/SQR[GARCH](1)			0.0007
			(0.0032)
EGARCH(1)			0.9957**
			(0.0009)
C	0.0034**	0.0034**	-0.0458**
	(0.0008)	(0.0008)	(0.0060)
Model Fit			
LL	-39845.55	-39844.81	-39937.95
Adjusted LL	33535.45	33536.20	33443.05
AIC	2.6527	2.6528	2.6589
Adjusted AIC	-2.2322	-2.2321	-2.2259
SIC	2.6539	2.6541	2.6603
Adjusted SIC	-2.2311	-2.2308	-2.2246

Table 4.16: Comparison of Modelling Approaches

This table reports the parameter estimates of the best performing models for each of the approaches denoted Tl – T13 in the text. The statistics shown are the log likelihood (LL), the Akaike Information Criterion (AIC), and the Schwarz Information Criterion (SIC). Column 2 presents the best performing GARCH model specification under each individual approach. Ranking of overall best approach according to AIC and SIC statistics is shown under the reported statistics in columns 4 and 5 respectively.

	Approach	Best GARCH Model Specification	LL	AIC (Rank)	SIC (Rank)
T1	Non-periodic GARCH model	EGARCH	31981.56	-2.1287 (13)	-2.1273 (13)
T2	Jointly estimated full dummy version of the PGARCH model	EGARCH	33788.20	-2.2483 (1)	-2.2438 (1)
T3	Two-step full dummy version of the PGARCH model	GARCH	33308.53	-2.2171 (8)	-2.2159 (8)
T4	Jointly estimated partial dummy version of the PGARCH model	EGARCH	33581.42	-2.2350 (4)	-2.2325 (4)
T5	Two-step partial dummy version of the PGARCH model	GARCH	33392.07	-2.2226 (7)	-2.2215 (7)
T6	Jointly estimated FFF version of the PGARCH model	EGARCH	33012.92	-2.1971 (9)	-2.1946 (9)
T7	Two-step FFF version of the PGARCH model	TGARCH	32670.37	-2.1745 (10)	-2.1731 (10)
T8	Jointly estimated Augmented FFF version of the PGARCH model	EGARCH	32652.07	-2.1731 (11)	-2.1706 (11)
T9	Two-step Augmented FFF version of the PGARCH model	TGARCH	32564.72	-2.1675 (12)	-2.1661 (12)

Table 4.16: Comparison of Modeling Approaches (continued)

This table reports the parameter estimates of the best performing models for each of the approaches denoted T1 – T13 in the text. The statistics shown are the log likelihood (LL), the Akaike Information Criterion (AIC), and the Schwarz Information Criterion (SIC). Column 2 presents the best performing GARCH model specification under each individual approach. Ranking of overall best approach according to AIC and SIC statistics is shown under the reported statistics in columns 4 and 5 respectively.

	Approach	Best GARCH Model Specification	LL	AIC (Rank)	SIC (Rank)
T10	Jointly estimated Spline version of the PGARCH model	EGARCH	33698.65	-2.2422 (3)	-2.2375 (3)
T11	Two-step Spline version of the PGARCH model	TGARCH	33551.49	-2.2332 (5)	-2.2318 (5)
T12	Jointly estimated Augmented Spline version of the PGARCH model	EGARCH	33736.84	-2.2447 (2)	-2.2399 (2)
T13	Two-step Augmented Spline version of the PGARCH model	TGARCH	33536.20	-2.2321 (6)	-2.2308 (6)
	Jointly estimated full dummy version of the PGARCH model allowing α to vary periodically	GARCH	33547.40	-2.2316	-2.2244

Figure 4.1: KLCI Returns

This figure shows the plot of the KLCI 5-minute returns for the 406 trading days used as the sample in the study. Index returns are computed by taking the first difference in log prices during various five-minute intervals over the trading days.

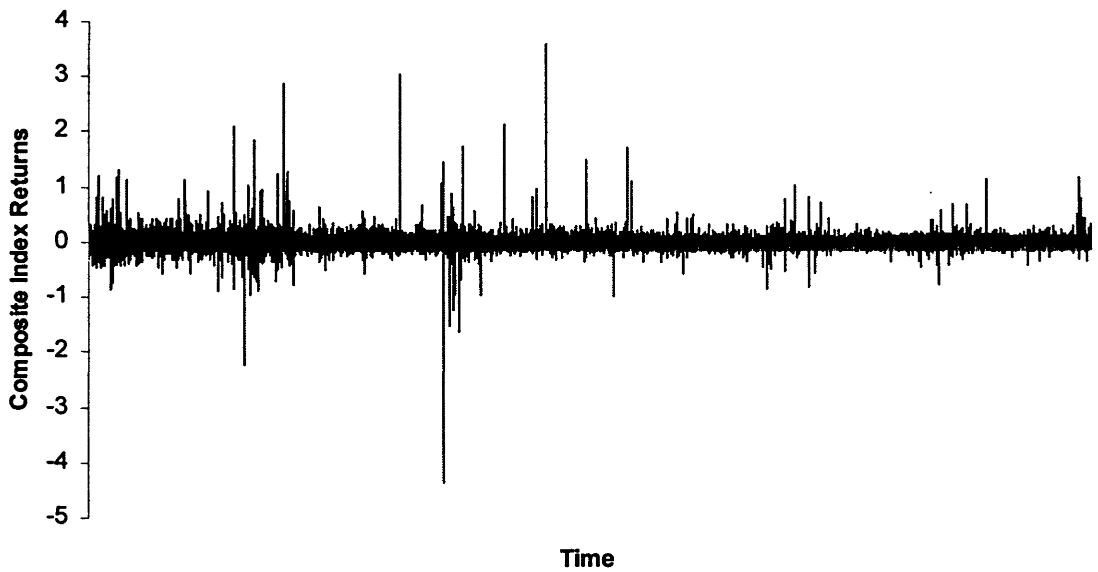


Figure 4.2: Intraday Periodicity

This figure shows the intraday volatility of the KLCI returns. Intraday volatility is computed by taking the means of absolute returns during various five-minute intervals over the trading day. The break in the curve indicates closure of trading during lunch hours.

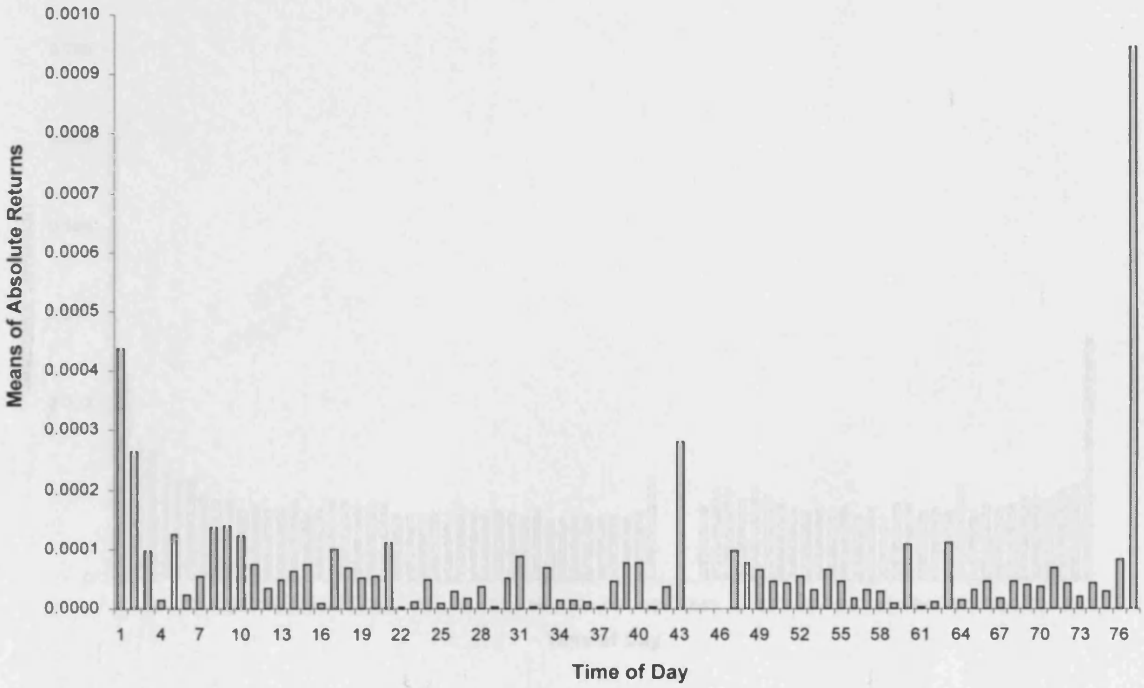


Figure 4.3: Intraday Periodicity

This figure shows the intraday volatility of the KLCI returns. Intraday volatility is calculated by taking the standard deviation of returns during various five-minute intervals over the trading day. The break in the curve indicates closure of trading during lunch hours.

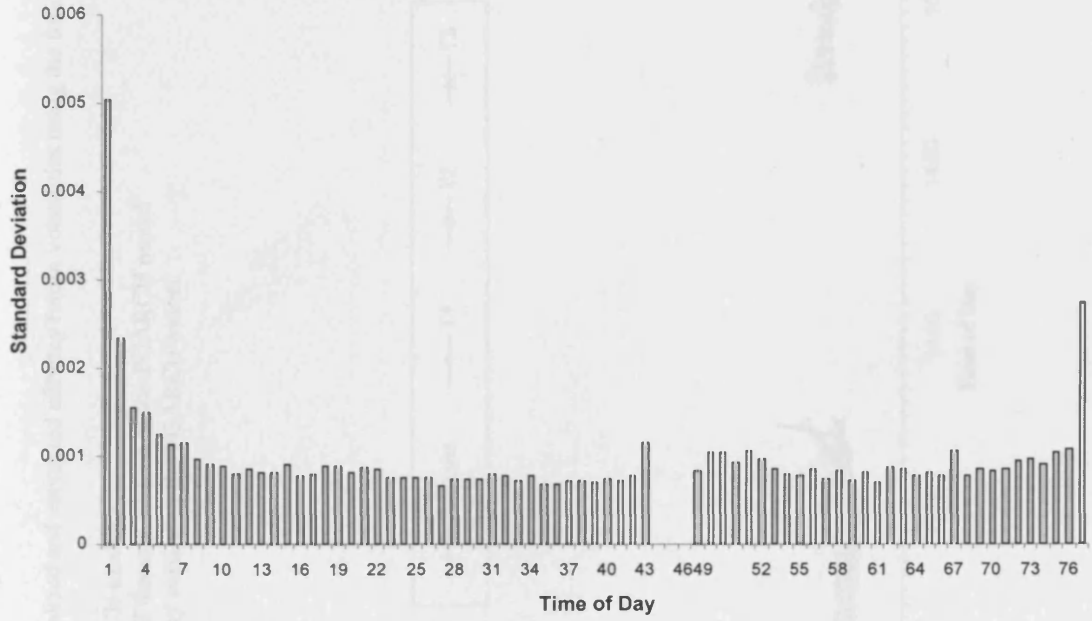


Figure 4.4: Intraday Return Volatility

The chart below compares the mean realized and estimated intraday return volatilities using the following approaches:

T1 = Non-periodic GARCH model
T2 = Jointly estimated full dummy version of the PGARCH model
T3 = Two-step full dummy version of the PGARCH model

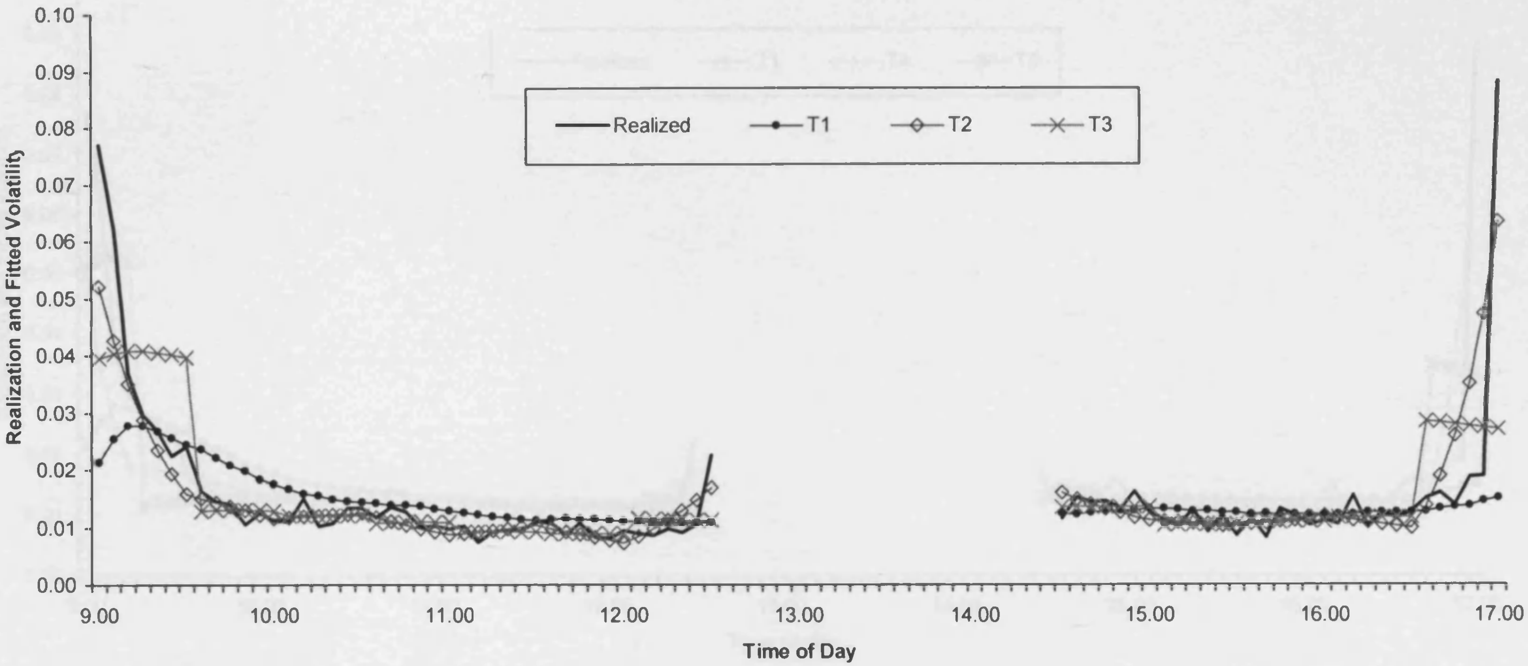


Figure 4.5: Intraday Return Volatility

The chart below compares the mean realized and estimated intraday return volatilities using the following approaches:

T1 = Non-periodic GARCH model
T4 = Jointly estimated partial dummy version of the PGARCH model
T5 = Two-step partial dummy version of the PGARCH model

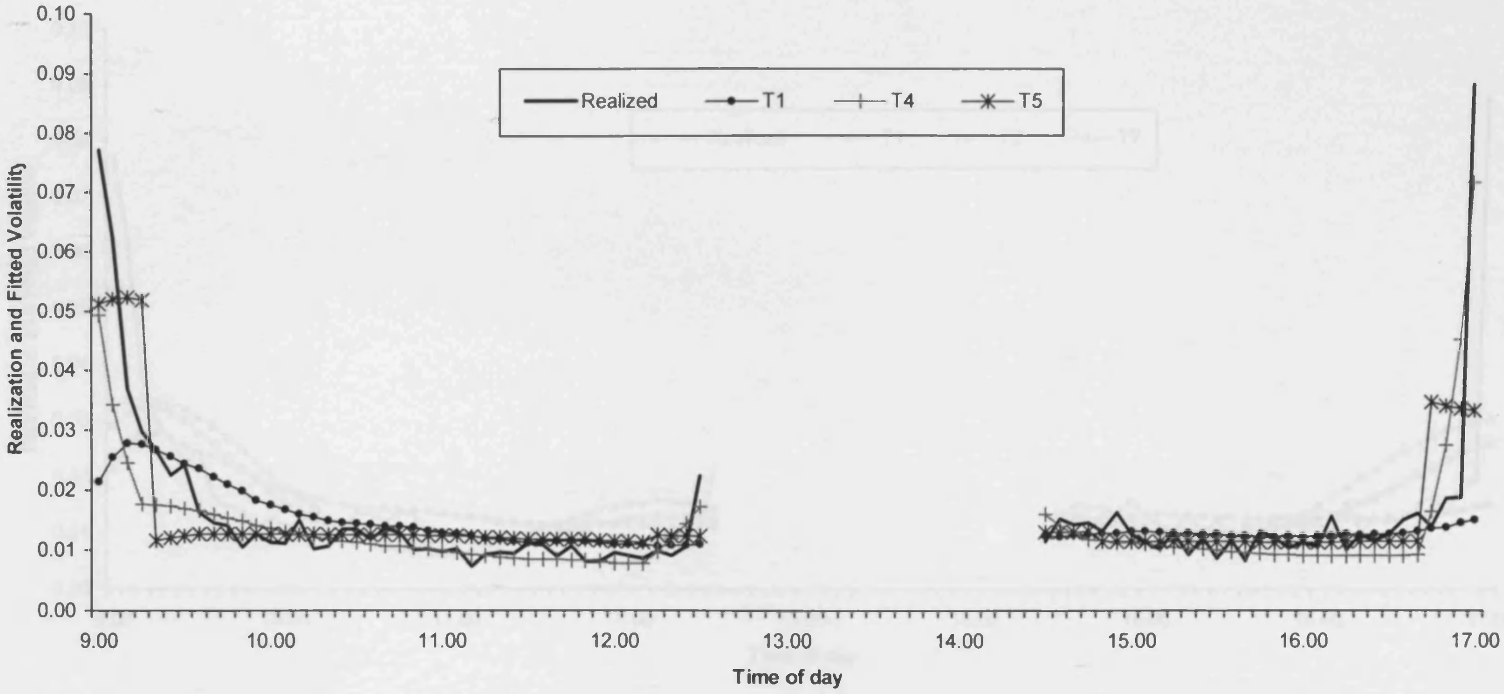


Figure 4.6: Intraday Return Volatility

The chart below compares the mean realized and estimated intraday return volatilities using the following approaches:

T1 = Non-periodic GARCH model
T6 = Jointly estimated FFF version of the PGARCH model
T7 = Two-step FFF version of the PGARCH model

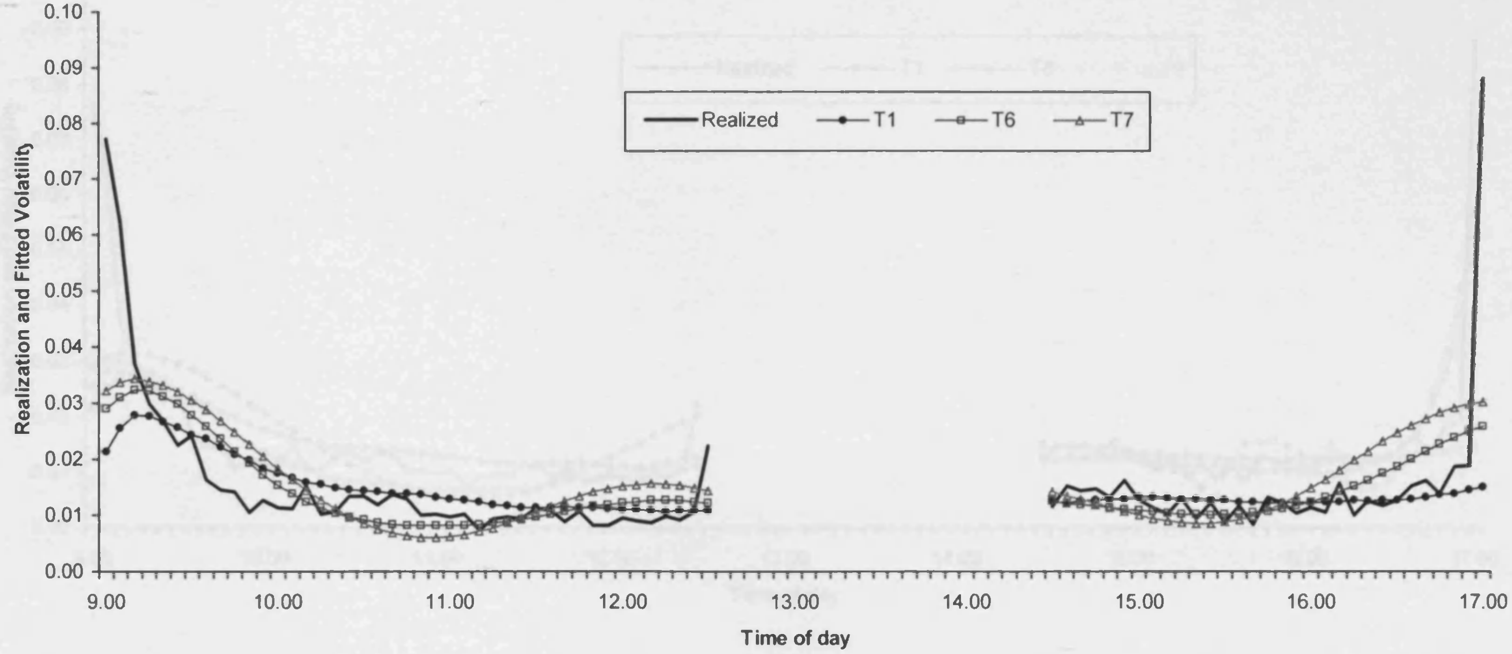


Figure 4.7: Intraday Return Volatility

The chart below compare the mean realized and estimated intraday return volatilities using the following approaches:

T1 = Non-periodic GARCH model
T8 = Jointly estimated Augmented FFF version of the PGARCH model
T9 = Two-step Augmented FFF version of the PGARCH model

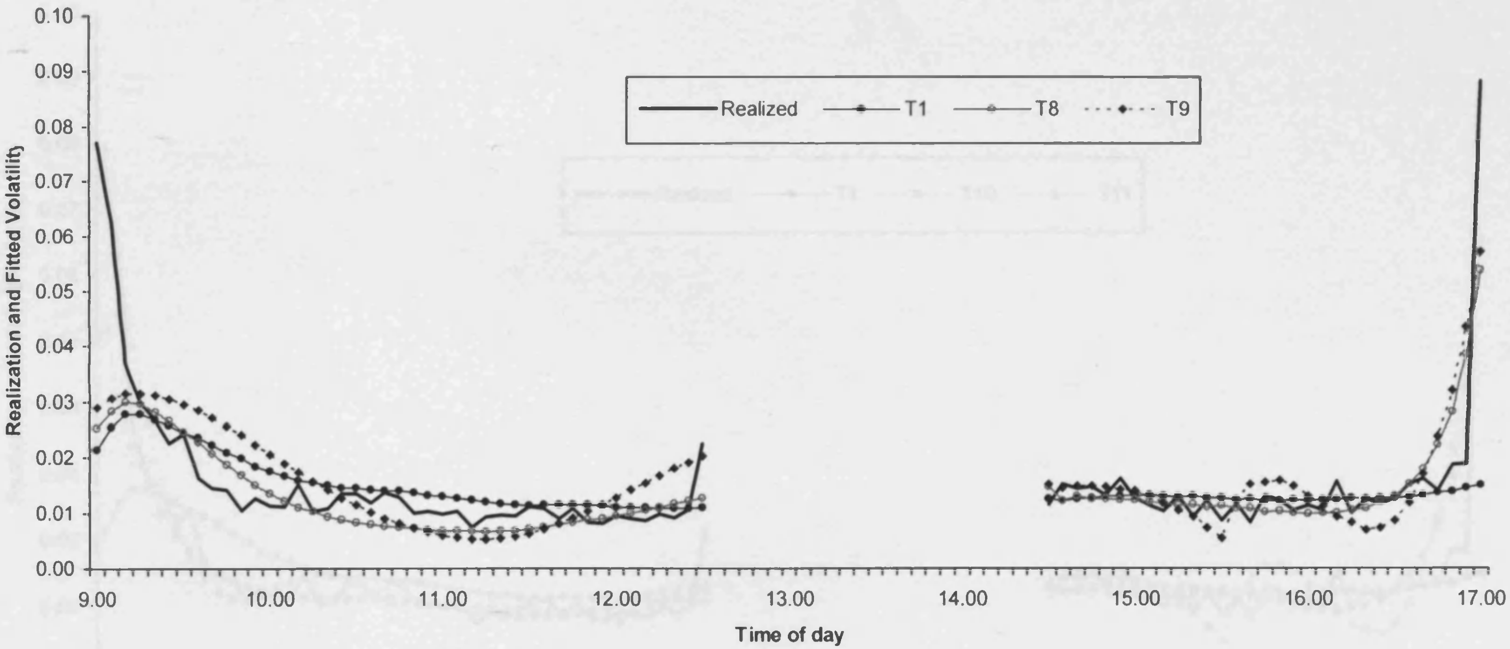


Figure 4.8: Intraday Return Volatility

The chart below compare the mean realized and estimated intraday return volatilities using the following approaches:

T1 = Non-periodic GARCH model
T10 = Jointly estimated Spline version of the PGARCH model
T11 = Two-step Spline version of the PGARCH model

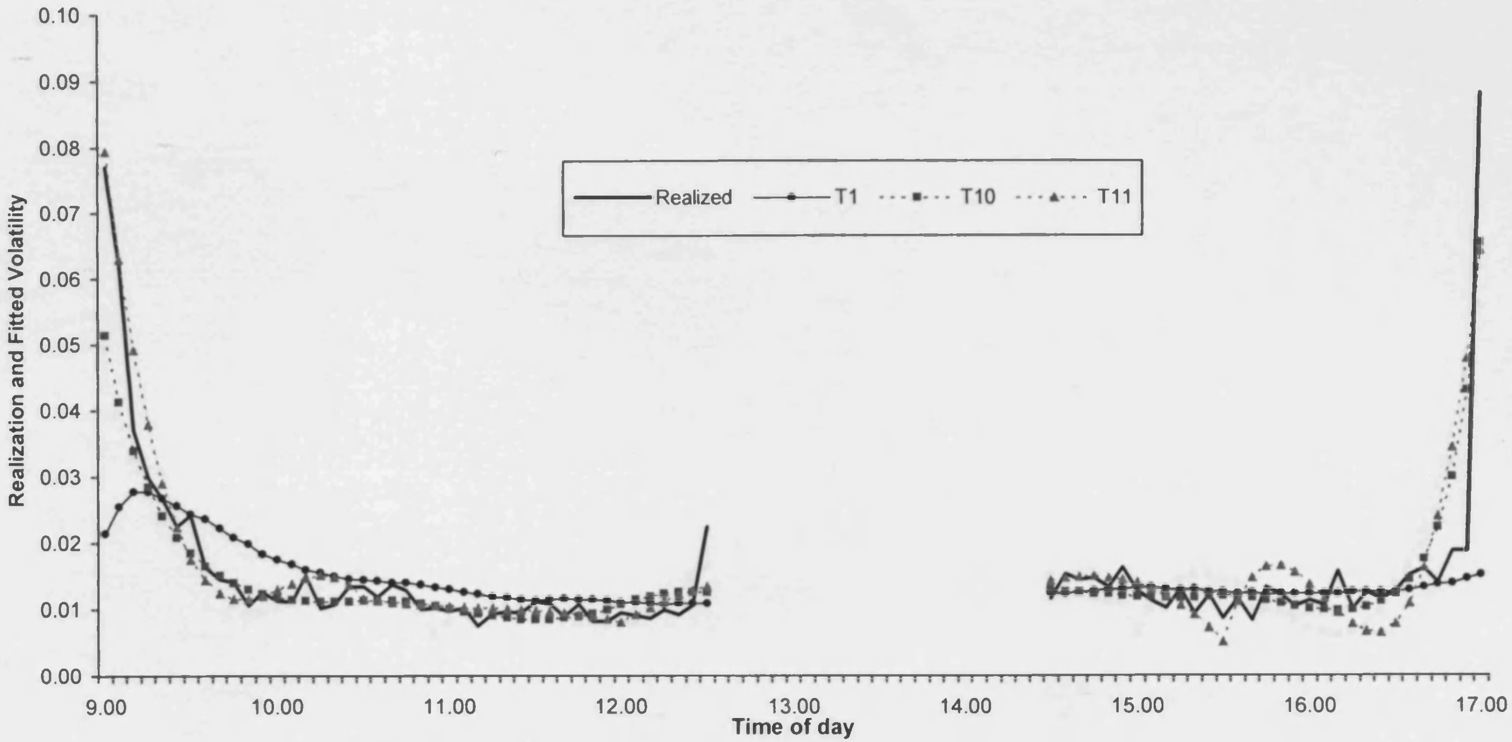


Figure 4.9: Intraday Return Volatility

The chart below compare the mean realized and estimated intraday return volatilities using the following approaches:

T1 = Non-periodic GARCH model
T12 = Jointly estimated Augmented Spline version of the PGARCH model
T13 = Two-step Augmented Spline version of the PGARCH model

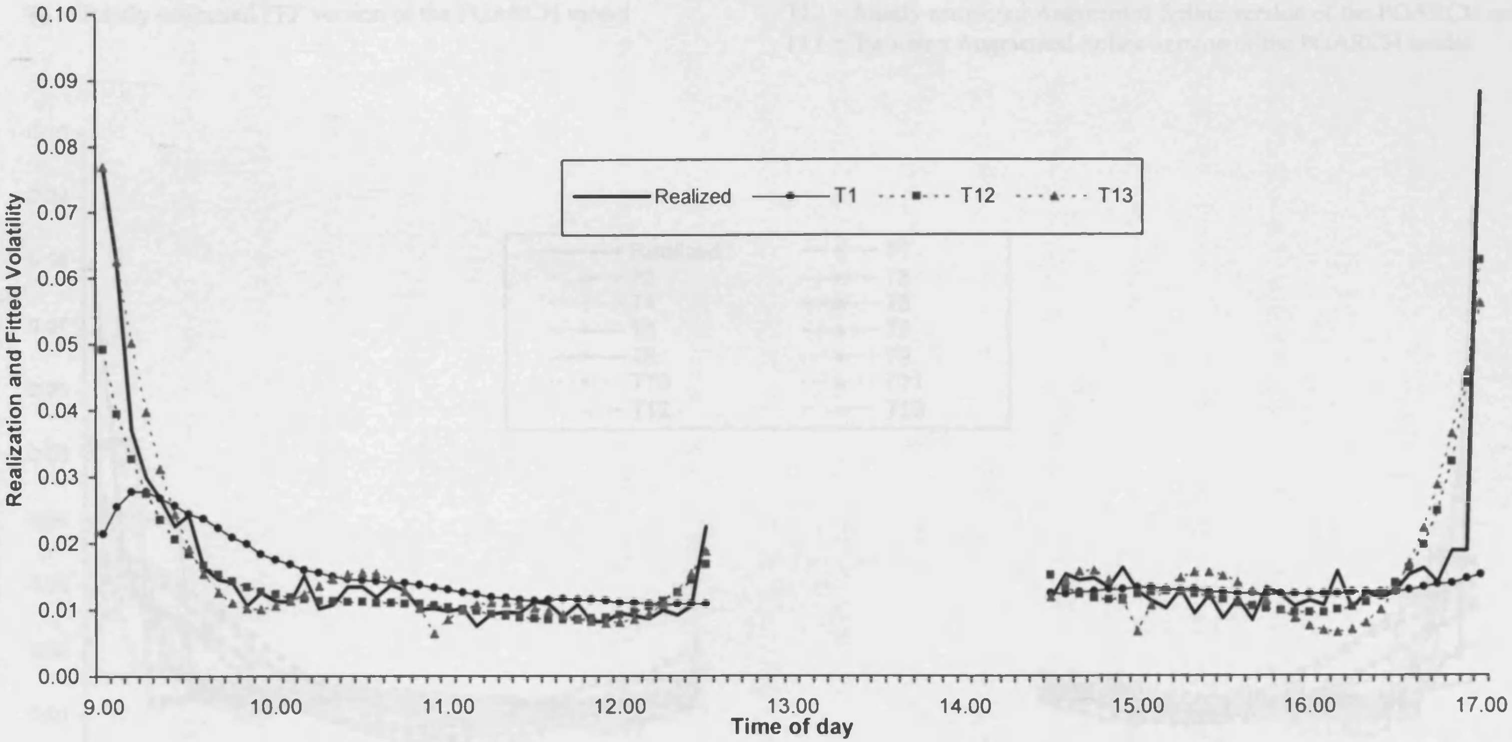
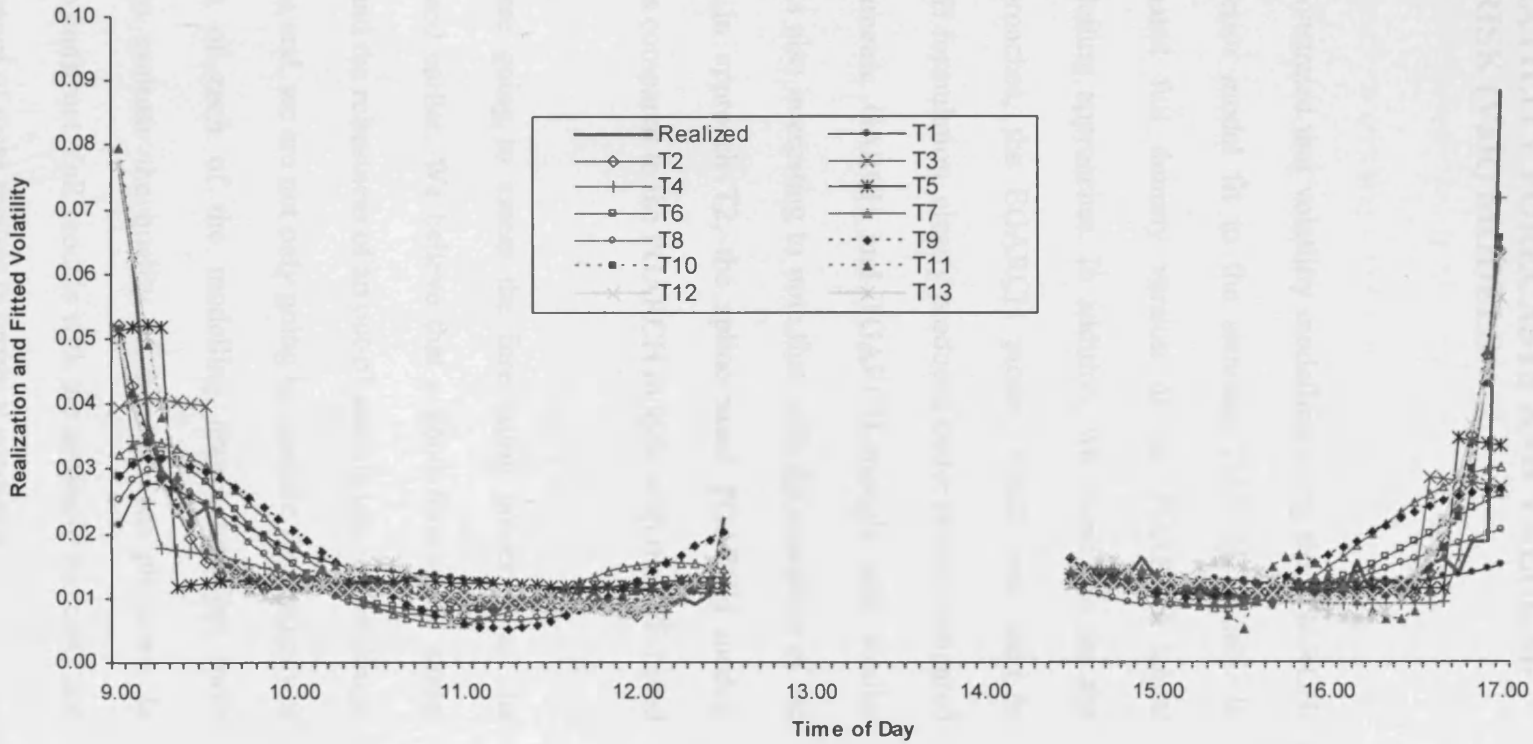


Figure 4.10: Intraday Return Volatility

This figure shows the mean realized and estimated intraday return volatilities using the following approaches:

T1 = Non-periodic GARCH model
T2 = Jointly estimated full dummy version of the PGARCH model
T3 = Two-step full dummy version of the PGARCH model
T4 = Jointly estimated partial dummy version of the PGARCH model
T5 = Two-step partial dummy version of the PGARCH model
T6 = Jointly estimated FFF version of the PGARCH model
T7 = Two-step FFF version of the PGARCH model
T8 = Jointly estimated Augmented FFF version of the PGARCH model
T9 = Two-step Augmented FFF version of the PGARCH model
T10 = Jointly estimated Spline version of the PGARCH model
T11 = Two-step Spline version of the PGARCH model
T12 = Jointly estimated Augmented Spline version of the PGARCH model
T13 = Two-step Augmented Spline version of the PGARCH model



CHAPTER 5

EVALUATING VOLATILITY FORECASTS AND VALUE-at-RISK (VaR) MODELS

5.0 Introduction

In Chapter 4, we demonstrated that volatility modelling using the PGARCH-based models produced superior model fit to the standard GARCH models. In particular, the jointly estimated full dummy version of the PGARCH model dominated all the other modelling approaches. In addition, we found that for the jointly estimated based approaches, the EGARCH model, which was used in conjunction with the PGARCH formulation, clearly produced better results compared with the results of the symmetric GARCH and TGARCH models with similar formulation. Furthermore, it is also interesting to note that with the exception of the dummy-based variable used in approach T2, the spline-based PGARCH models showed superior performances compared to the PGARCH models with the FFF-based variables.

In this chapter, we are going to assess the forecasting power of all the modelling approaches discussed earlier. We believe that a good forecasting model should be one that can withstand the robustness of an out-of-sample test - a test design that is closer to reality. To this end, we are not only going to measure the accuracy of the forecasting performance of each of the modelling approaches, but more importantly, we are going to evaluate the quality of the forecasts produced. In addition, we are also going to construct VaR models with the available forecasts and evaluate the economic significance of these models in terms of adequacy.

Following this introduction, section 5.1 discusses some aspects of volatility forecasting using high-frequency data, as well as some details on the VaR measures. Section 5.2 details the construction of data that provide the basis for our subsequent empirical analysis. We will also discuss the details of how the volatility forecasts generated by the thirteen volatility modelling approaches are evaluated in terms of performance and accuracy. We then discuss how the VaR models are constructed and the tests applied to evaluate the quality of these VaR models. In section 5.3, we discuss the results of the in-sample fits and the out-of-sample forecasting performance of the various GARCH models estimated using the thirteen volatility modelling approaches described in Chapter 4. We also compare and evaluate the VaR models constructed from the various GARCH forecasts using the coverage tests of Christoffersen (1998) and the regression-based tests developed by Clements and Taylor (2003). We finish in section 5.4 with concluding remarks. All results are reported at the end of the chapter.

5.1 Chapter Background

There is no question that the accuracy of volatility forecasting has received much attention recently, not only from academics and financial market participants, but also from policy makers who are concerned about the stability and the well being of the economy. The growing attention on volatility forecasting is not surprising considering the impact of asset return volatility on financial markets. However, the complex dynamics inherent in the volatility process mean that rather different results may be obtained depending on the model used and on the market conditions. For example, large price swings in stock, bond, foreign exchange, commodity and energy markets in today's globally interconnected financial markets have proven to be a serious concern to all parties. Numerous instances of financial instability, such as the 1997 Asian financial crisis (with the associated collapse of stock markets), and a number of widely publicized losses suffered by major corporations and banks, are partly due to adverse and volatile market movements.[1] For these reasons, much effort has been made in trying to understand the dynamics of asset return volatility. Only when the nature of return volatility is sufficiently understood can the search for superior modelling techniques begin.

This intensive endeavour, pursued by academics and practitioners alike, has resulted in a large literature on volatility modelling and forecasting.[2] However, most of the works done in this area have until recently utilised interday data rather than intraday data. The development of intraday databases, spanning a host of financial instruments and markets, has presented a new challenge to the modelling and forecasting of asset return volatility. The availability of the so-called "ultra high frequency" (Engle, 2000) tick data has made it possible to use high frequency data to model the intraday volatility patterns of asset returns. This application of volatility models within the day is a natural extension of the daily models examined so widely in the past. The outcomes of many of the studies have resulted in a richer understanding of the intraday volatility generating process, revealing some behaviours that were not observed at daily and lower frequencies.[3] The obvious question now is "does the use of high frequency data produce better volatility forecast performance?" A study by Andersen, Bollerslev, Diebold and Labys (2003), for example, found that this could be the case. They go further by suggesting that forecast improvement is not only due to the application of high frequency data, but also that the information

[1] See, for example, the collapses of the Enron Corporation of the United States in 2001, and Barings Bank of the UK in 1995.
[2] See, for example, Poon and Granger (2003), for a detailed survey of 93 published and working papers.
[3] See, for example, Taylor (2004), on intraday and interday volatility periodicities in cocoa futures contracts traded on the *Euronext.liffe* exchange.

content in high frequency data is useful for forecasting at longer horizons, such as monthly or quarterly periods. On the negative side, however, they found that standard volatility models used for forecasting at the daily level cannot readily accommodate the information in intraday data. Moreover, the results show that when these models are specified for the intraday data, they are generally found not to be sufficiently successful in capturing the longer interdaily volatility movements. This finding does not augur well, for it highlights a serious limitation of the GARCH models in modelling intraday asset return volatility. This limitation exists because the GARCH models were originally developed and designed to capture features of financial time series measured at daily (and lower) frequencies. An earlier study by Andersen and Bollerslev (1998a) confirmed that this is the case when they found that the GARCH models provide seemingly poor forecasts of daily volatility when standard forecast evaluation criteria are imposed on high frequency data. Predictably, these findings have led to the perception that GARCH models are of limited practical use in studies involving the use of high frequency data.

What could be the shortcomings of the GARCH models when applied to high frequency data? Empirical findings by Figlewski (1997) suggest that all GARCH models share two significant weaknesses as forecasting tools. Firstly, the models seem to require a large number of observations or data points for the estimation to be robust. Secondly, the GARCH models are subject to the general problem of fitting the sample data. It is found that, in general, the more complex the construction of the GARCH model (i.e. the presence of more parameters), the better it will tend to fit a given sample data and the quicker it will tend to fail out-of-sample. For any procedure to be useful in forecasting, it must be sufficiently stable over time that one can fit coefficient estimates on historical data and be reasonably confident that the model will continue to hold over time. Nevertheless, the standard GARCH model and its derivatives are still useful because they serve as a benchmark for comparison and evaluation purposes with respect to other new models such as the stochastic volatility models and models based upon "realized volatility" proposed by Andersen, Bollerslev, Diebold and Labys (2003).

As mentioned above, one important use of volatility forecasts, which has grown substantially in importance over recent years, is as an input to financial risk management. One of the most popular approaches to financial risk management is what is known as VaR. In general, VaR is a measure of the market risk of a portfolio. It quantifies, in monetary terms, the exposure of a portfolio to future market fluctuations.[4] It is, at present, a regulatory disclosure requirement by the Basle Committee on Banking Supervision (1996, 1998) for capital adequacy and in financial reports. While the concept is simple and attractive, there is no consensus on how best to implement VaR. Rather, in practice, there are a wide variety of alternative models that are used in the generation of VaR forecasts, with each alternative model tending to yield different VaR forecasts. As such, there is a need to assess the quality of these VaR forecasts, for it is often found that different methodologies can yield different VaR measures for the same portfolio, sometimes leading to significant errors in risk management. Consequently, there has been a surge in interest in the empirical literature in measuring the quality of alternative VaR implementations and in tackling the problem of model selection. Among the more popular volatility models that are used in VaR estimation are the GARCH-based models. They have been tested

[4] J.P Morgan (1996), for example, defines VaR as a measure of the maximum potential change expected in the value of a portfolio with a given probability over a pre-determined horizon.

extensively and have been proven useful in the generation of VaR forecasts in developed markets, but less so in emerging capital markets.[5]

The aim of this chapter is therefore to combine and advance the two literatures in volatility forecasting and financial risk management via VaR modelling. This chapter will attempt to address the question of the usefulness of GARCH-based models in explaining past volatility and forecasting the future volatility of stock index returns for the KLSE. In this respect, we undertake an extensive analysis of out-of-sample forecasts in an effort to gain a better understanding of volatility predictability. To this end, we employ the GARCH-based models and the thirteen volatility modelling approaches introduced in the previous chapter. More specifically, we will evaluate whether the incorporation of periodic components in the volatility models produces better forecasting performance than the standard non-periodic GARCH model. We will address whether one volatility modelling approach is significantly better than another among the thirteen approaches employed in the earlier chapter. To do this, we will apply the forecast comparison test proposed by Diebold and Mariano (1995), and the forecast encompassing test developed by Harvey, Leybourne and Newbold (1998). These two tests are important because they provide statistical significance to the quality of forecasts produced by the various modelling approaches. In view of the economic significance of VaR, we will also assess the adequacy and quality of the VaR measures produced through the forecasts generated by the GARCH-based models by applying the coverage tests of Christoffersen (1998) and the regression-based tests proposed by Clements and Taylor (2003).

As mentioned earlier, this study is important because the KLSE is one of the largest emerging capital markets in Asia. It has different risk and return characteristics

[5] See, for example, Christoffersen, Diebold and Schuermann (1998), Lopez (1999), Berkowitz and O'Brien (2002), Brooks and Persand (2003), Giot and Laurent (2004), and Bredin and Hyde (2004).

as well as different institutional structures from developed markets. It is hoped that this study will cast light on the behaviour and intraday forcastability of volatility via an evaluation of VaR risk management on the KLSE. We also hope to contribute to the scarce literature on similar studies in emerging capital markets.

5.2 Data and Methodology

The same set of KLCI data described in Chapter 4 is utilized in this chapter. The data set consists of high frequency data measured at 5-minute intervals over the period from 29 January 2001 to 29 December 2002. This set of data produces 30044 observations or 406 days with 72 5-minute frequency observations in a trading day. This setting will be the benchmark for testing a number of related objectives. The basic methodology involves the estimation of the various GARCH-based model parameters using an initial set of data and the application of these parameters to later data, thus forming out-of-sample forecasts. The total sample of 30044 observations is therefore split into two parts: the first 22644 observations (306 days) are used for the in-sample estimation of the parameters of the various GARCH-based models employed by the thirteen volatility modelling approaches. The last 7400 observations are then used to generate 5-minute one-step-ahead out-of-sample forecasts for all the available GARCH-based models used in the study for the 100-day forecast period.[6]

It has been established from the findings from the previous chapter that the KLCI return volatility is periodic. It has also been found from Chapter 4 that modelling approaches that incorporate the periodic components in the conditional volatility process exhibit superior model fit over the standard non-periodic GARCH model approach. However, we believe that the success of a volatility modelling

[6] Other forecast periods were also considered, with similar results.

approach lies in its out-of-sample forecasting power. Therefore, it is of interest to compare the forecasting performance of the PGARCH-based models with those of the non-periodic GARCH model, i.e., the standard GARCH model without periodic components. It would also be insightful if we could establish whether the GARCH-based models (periodic and non-periodic) are superior in terms of forecasting accuracy to the naive model, which will be based on the estimate of historical volatility over the estimation period. We compute the historical variance over the estimation period as:

$$\sigma_t^2 = \frac{1}{T}\sum_{t=1}^{T} r_t^2 \quad . \tag{5.1}$$

where $r_t = R_t - \mu$; R_t is the 5-minute interval returns, μ is the expected return, and T is the total number of 5-minute intervals observed in the estimation period. For convenience, we reproduce the thirteen volatility modelling approaches and the best performing GARCH-based model for each individual approach in Table 5.1 below.

The first phase of this chapter is to evaluate the forecasting performance of the GARCH-based models for the thirteen approaches described in Table 5.1. We begin by estimating (in-sample) the parameters of the GARCH-based models for the T1 to T11 modelling approaches. The in-sample period covers 306 days or 22644 5-minute observations. All GARCH models are estimated by maximum likelihood with Bollerslev-Wooldridge robust QML covariance/standard errors. In addition, as before, the fit of the models is measured by the AIC and the SIC statistics as well as the LL function. We then rank each approach according to the model fit. The best approach is ascertained by the model that gives the minimum AIC and SIC values. When the AIC and SIC values are in conflict, the best model is ranked by the SIC measure, as the SIC embodies a much stiffer penalty term than AIC, and therefore is preferred in a forecast evaluation setting. The next best modelling approach is then ranked accordingly based on the next GARCH-based model that produces the minimum AIC and SIC values. The process continues until all the approaches are appropriately ranked. The second phase involves generating 5-minute one-step-ahead forecasts for the 100-day out-of-sample period for all the available models and modelling approaches, using the in-sample parameters estimated previously. We therefore have thirteen sets of 5-minute one-step-ahead forecasts, which will then be evaluated in terms of performance and accuracy.

Table 5.1: Comparison of Modelling Approaches

This table shows the thirteen volatility modelling approaches used in Chapter 4. The best performing GARCH-based model in terms of model fit is given in the second column.

	Approach	Best GARCH Model Specification
T1	Non-periodic GARCH model	EGARCH
T2	Jointly estimated full dummy version of the PGARCH model	EGARCH
T3	Two-step full dummy version of the PGARCH model	GARCH
T4	Jointly estimated partial dummy version of the PGARCH model	EGARCH
T5	Two-step partial dummy version of the PGARCH model	GARCH
T6	Jointly estimated FFF version of the PGARCH model	EGARCH
T7	Two-step FFF version of the PGARCH model	TGARCH
T8	Jointly estimated Augmented FFF version of the PGARCH model	EGARCH
T9	Two-step Augmented FFF version of the PGARCH model	TGARCH
T10	Jointly estimated Spline version of the PGARCH model	EGARCH
T11	Two-step Spline version of the PGARCH model	TGARCH
T12	Jointly estimated Augmented Spline version of the PGARCH model	EGARCH
T13	Two-step Augmented Spline version of the PGARCH model	TGARCH

5.2.1 Evaluating Volatility Forecasts

There are a variety of statistics available to evaluate and compare the performance of forecasts produced by volatility models.[7] In this study, two measures are used to evaluate the accuracy of forecasts: viz., the mean squared forecast error (MSFE) and the mean absolute forecast error (MAFE). These measures are defined as:

$$\text{MSFE} = \frac{1}{T_2}\sum_{t=1}^{T_2}(y_{t+h} - \hat{y}_{t+h})^2 .$$

$$\text{MSFE} = \frac{1}{T_2}\sum_{t=1}(y_{t+h} - \hat{y}_{t+h})^2 . \tag{5.2}$$

$$\text{MAFE} = \frac{1}{T_2}\sum_{t=1}^{T_2}\left|y_{t+h} - \hat{y}_{t+h}\right| . \tag{5.3}$$

where y_{t+h} is the realization of the series at time $t+h$, $\hat{y}_{t+h}$ is the *h*-step ahead forecast of the series using data observed up to and including time *t*, and T_2 is the number of *h*-step-ahead forecasts considered. The MSFE provides a quadratic loss function which disproportionately weights large forecast errors more heavily relative to the MAFE and thus may be useful in forecasting situations when large forecast errors are disproportionately more serious than small errors.

We compute the MSFE and the MAFE for each set of one-step-ahead forecasts generated. We then rank each of the thirteen approaches according to which produces the most accurate forecasts. Starting with the MSFE, we first rank the approach that gives the smallest MSFE statistic. Next, we take the second smallest MSFE statistics produced by the next relevant approach. The process continues until all approaches are ranked accordingly. We then repeat the same process with the MAFE statistics for all the volatility modelling approaches. At the end of the exercise, we have two sets of forecasting performance rankings for T1 to T13. Finally, we

[7] For a detailed survey on the popular evaluation measures used in the literature, please refer to Poon and Granger (2003).

compute the MSFE and MAFE for the unconditional variance estimator for comparison purposes.

To compare the predictive accuracy of alternative forecasts, we employ an asymptotic test of the null hypothesis of no difference in the accuracy of two competing forecasts proposed by Diebold and Mariano (1995). This is a convenient test of the null hypothesis that the forecasts from two models do not differ significantly. We assume that the time t loss associated with a forecast i is an arbitrary function of the realization and prediction, $g(y_t, \hat{y}_{it})$; specifically, it is assumed to be a direct function of the forecast error, $g(y_t, \hat{y}_{it}) = g(e_{it})$. Under this assumption, the null hypothesis of equal forecast accuracy for two competing forecasts is $\mathrm{E}[g(e_{it})] = \mathrm{E}\left[g(e_{jt})\right]$ or $\mathrm{E}[d_t] = 0$, where $d_t \equiv \left[g(e_{it}) - g(e_{jt})\right]$ is the loss differential. Thus, the equal accuracy null hypothesis is equivalent to the null hypothesis that the population mean of the loss-differential series is zero.

Let $$\bar{d} = \frac{1}{T_2}\sum_{t=1}^{T_2}\left[g(e_{it}) - g(e_{jt})\right], \tag{5.4}$$

denote the sample mean loss differential over T_2 forecasts, and let $g(e_{it})$ be a general function of forecast errors (e.g. MAFE or MSFE); then the Diebold Mariano (1995) test statistic (henceforth denoted DM) is given by:

$$DM = \frac{\bar{d}}{\sqrt{\frac{2\pi\hat{f}_d(0)}{T_2}}}. \tag{5.5}$$

where $\hat{f}_d(0)$ is a consistent estimate of the spectral density of the loss differential at frequency zero. The DM statistic has a standard normal distribution with mean zero and unit variance under the null hypothesis. The loss functions adopted in this study are the squared function (MSFE) and the absolute function (MAFE). We proceed by comparing the MSFE and MAFE values associated with the forecasts of each of the thirteen approaches, T1 to T13, with the naive model, denoted by M1. Next, we apply the DM asymptotic test for each competing pair of forecasts among the thirteen volatility modelling approaches. This exercise produces 78 forecast quality comparisons.

We then apply the test of forecast encompassing developed by Harvey, Leybourne and Newbold (1998). Forecast encompassing refers to whether or not the forecasts from a competing model contain information that is missing from the forecasts of the original model. If they do not, then the forecasts from the competing model are said to be 'encompassed' by the forecasts from the original model. Furthermore, a forecast is considered 'conditionally efficient' if the variance of the forecast error from a combination of that forecast and a competing forecast are equal to or greater than the variance of the original forecast error. Therefore, a forecast that is conditionally efficient 'encompasses' the competing forecast. Harvey, Leybourne and Newbold (1998) developed an encompassing test based on the fact that if the forecasts from model 1 encompass the forecasts from model 2, then the covariance between e_{1t} and $e_{1t} - e_{2t}$ will be negative or zero (e_{1t} and e_{2t} are the two sets of forecast errors from model 1 and model 2 respectively). The alternative hypothesis is that the forecasts from model 1 do not encompass those from model 2, in which case the covariance between e_{1t} and $e_{1t} - e_{2t}$ will be positive. The Harvey, Leybourne and Newbold (1998) test statistic (henceforth denoted HLN) is formulated as follows:

$$\text{HLN} = \frac{\bar{c}}{\sqrt{\text{var}(\bar{c})}}. \tag{5.6}$$

where $\bar{c} = \frac{1}{T_2}\sum_{t=1}^{T_2} c_t$, $c_t = e_{1t}(e_{1t} - e_{2t})$, and T_2 is the number of observed forecasts. The HLN statistic has an asymptotic standard normal distribution. Similar to the steps applied in the DM asymptotic test above, we compare the unconditional variance estimator with each forecast of the individual volatility modelling approaches, T1 to T13. Finally, we proceed by comparing the forecasts for each pair of competing volatility modelling approaches. This again results in 78 forecast quality comparisons.

5.2.2 Evaluation of VaR Models

To examine the economic significance of the volatility forecasts, we evaluate the quality of the VaR models constructed from the available GARCH-based models forecasts for each of the thirteen volatility modelling approaches. In addition to these approaches, we also consider the *RiskMetrics* modelling approach, a methodology for measuring market risk, which is widely used in the financial industry.[8] The VaR measures associated with these models are constructed using the 5-minute one-step-ahead forecasts with a 100-day forecast horizon and both 95% and 99% confidence levels, as recommended by the Basle Committee. Assuming that the return series is conditionally normally distributed, the VaR measures are computed by multiplying the conditional standard deviation $(\hat{\sigma}_t)$ by the appropriate percentile point on the normal distribution ($\alpha = 0.05$ and $\alpha = 0.01$ respectively). With this, we construct the dynamic interval forecasts (i.e., VaR measure) using the out-of-sample forecasts

[8] The RiskMetrics VaR model adopts an exponentially weighted moving average approach to return volatility modelling. In particular, it assumes that return volatility evolves as follows:

$$h_t = (1-\beta)\varepsilon_{t-1}^2 + \beta h_{t-1},$$

where $\varepsilon_t = R_t - \mu$ and β is given a value between 0.94 and 0.99, and in this study, we assume that $\beta = 0.94$.

produced above, which are designed to cover 95% and 99% future outcomes. Formally, the lower and upper limits of the intervals are computed as:

$$\left\{ \frac{-\alpha \hat{\sigma}_t}{2}, \frac{\alpha \hat{\sigma}_t}{2} \right\}, \tag{5.7}$$

for an interval with a nominal α x 100% coverage.

For comparison purposes, in view of the presence of fat tails in the distribution of the KLCI returns series, we also consider VaR measures that are computed by multiplying the conditional standard deviation $(\hat{\sigma}_t)$ by the appropriate percentile point on the Student t distribution ($\alpha = 0.05$ and $\alpha = 0.01$ respectively) to control for the effect of the fat tails. In the current application, we use Student t distributions with 4 and 24 degrees of freedom to approximate the distribution of KLCI returns series.

In order to evaluate the quality and adequacy of the various VaR measures, we adopt the framework for interval forecast evaluation developed by Christoffersen (1998). Christoffersen emphasizes testing what is known as the "conditional coverage" of the interval forecasts. The importance of testing "conditional coverage" arises from the observation of volatility clustering in many financial time series. Good interval forecasts should be narrow in tranquil times and wide in volatile times, so that observations falling outside a forecasted interval are spread over the entire sample, and do not come in clusters. A poor interval forecast may produce correct unconditional coverage, yet it may fail to account for higher-order time dynamics. In this case, the symptom that would be observed is a clustering of failures.

Christoffersen (1998) begins by classifying an interval forecast as a success or a failure. This classification is achieved by means of a simple indicator function. The indicator function, I_t, takes values as follows:

$$I_t = 1 \quad \text{if } y_t \in \left[L_{t|t-1}(p), U_{t|t-1}(p)\right], \tag{5.8}$$

$$= 0 \quad \text{otherwise.}$$

where y_t denotes a sequence of time series observations, $L_{t|t-1}(p)$ denotes the lower level, and $U_{t|t-1}(p)$ the upper level of an out-of-sample interval forecast which nominally covers a proportion p of the possible outcomes, and is made at period *t-1* for the following period *t*. Christoffersen (1998) defines a set of *ex ante* interval forecasts as being efficient with respect to the information set (denoted Ω_{t-1}) if the conditional expectation of I_t equals p, that is, $\mathrm{E}\left[I_t \mid \Omega_{t-1}\right] = p$. If one restricts the information set to past values of the indicator function, $\Omega_t = \left\{I_t, I_{t-1}, \ldots\right\}$, then this is equivalent to saying that $\left\{I_t\right\}$ is independently and identically distributed (i.i.d) Bernoulli with parameter p. To test for the "correct conditional coverage", Christoffersen (1998) develops a three-step testing procedure: a test for "correct unconditional coverage", a test for "independence", and a test for "correct conditional coverage".

In the test for correct unconditional coverage, the null hypothesis of the failure probability α is tested against the alternative hypothesis that the failure probability is different from α, under the assumption of an independently distributed failure process. We follow the likelihood ratio framework used by Christoffersen (1998). We test for the "correct unconditional coverage" by evaluating for each I_t series obtained from the out-of-sample forecasts available with $\mathrm{E}\left[I_t \mid \Omega_{t-1}\right] = p$ for both 95% and 99% confidence levels. This is performed for VaR estimation based on out-of-sample forecasts with 100-day out-of-sample forecasts. We assess the models' performances by first computing the empirical failure rate for both the left and right tails of the distribution of returns. By definition, the failure rate is the number of times returns exceed the forecasted one-period-ahead VaR measure. If the VaR model is correctly specified, the failure rate should be equal to the pre-specified VaR level of $\alpha = 0.01$ and $\alpha = 0.05$, respectively. As the computation of the empirical failure rate defines a sequence of yes/no observations, it is possible to test the following hypothesis:

$$H_0 : \hat{f} = \alpha , \tag{5.9}$$

and the alternative hypothesis

$$H_1 : \hat{f} \neq \alpha , \tag{5.10}$$

where $\hat{f}$ is the observed failure rate.

Based on (5.8), we can reformulate hypotheses (5.9) and (5.10) with the following null hypotheses:

$$\mathrm{E}[I_t] = p , \tag{5.11}$$

against the alternative hypothesis:

$$\mathrm{E}[I_t] \neq p . \tag{5.12}$$

We define the likelihood function as:

$$L(\pi; I_1, I_2, ..., I_n) = (1-\pi)^{n_0} \pi^{n_1}, \qquad \pi \in \Pi = [0,1] \tag{5.13}$$

where $n_1 = \sum_{j=1}^{n} I_j$ is the number of ones in the indicator series; and $n_0 = n - n_1$ is the number of zeros in the indicator series.

We then apply the following likelihood ratio statistic to test for unconditional coverage:

$$LR_{UC} = -2\ln\{L(p) / L(\hat{\pi})\} \sim \chi^2_{(1)} \tag{5.14}$$

where $L(p) = (1-p)^{n-n_1} p^{n_1}$, and $L(\hat{\pi}) = (1-\hat{\pi})^{n-n_1} \hat{\pi}^{n_1}$.

The likelihood ratio statistic is the ratio of the likelihood under the null hypothesis to the likelihood evaluated under the maximum likelihood estimate (MLE) $\hat{\pi}$ over the entire parameter space Π. We define $\hat{\pi}$ as the observed sample proportion of "successes", where $\pi = n_1/(n_0 + n_1)$. Under the null hypothesis H_0, the likelihood ratio statistic has a chi-square distribution with 1 degree of freedom.

In the test for "independence", the hypothesis of an independently distributed failure process is tested against the alternative hypothesis of a first-order Markov failure process. The likelihood ratio statistic is computed as:

$$LR_{IND} = -2\ln\left[\frac{(1-\hat{\pi}_2)^{(n_{00}+n_{10})}\,\hat{\pi}_2^{(n_{01}+n_{11})}}{(1-\hat{\pi}_{01})^{n_{00}}\,\hat{\pi}_{01}^{n_{01}}\,(1-\hat{\pi}_{11})^{n_{10}}\,\hat{\pi}_{11}^{n_{11}}}\right] \sim \chi_1^2. \tag{5.15}$$

where n_{ij} is number of i values followed by a j value in the I_t series, $i, j = 0,1$, $\pi_{ij} = \Pr\{I_t = i \mid I_{t-1} = j\}$, $\hat{\pi}_{01} = n_{01}/(n_{00}+n_{01})$, $\hat{\pi}_{11} = n_{11}/(n_{10}+n_{11})$, and $\hat{\pi}_2 = (n_{01}+n_{11})/(n_{00}+n_{01}+n_{10}+n_{11})$.

Finally, the test of "correct conditional coverage" is achieved by testing the null hypothesis of an independent failure process with failure probability α against the alternative hypothesis of a first order-Markov failure process with a different transition probability matrix. The test for correct conditional coverage is calculated as:

$$LR_{CC} = -2\ln\left[\frac{p^{n_1}(1-p)^{n_0}}{(1-\hat{\pi}_{01})^{n_{00}}\,\hat{\pi}_{01}^{n_{01}}\,(1-\hat{\pi}_{11})^{n_{10}}\,\hat{\pi}_{11}^{n_{11}}}\right] \sim \chi_2^2. \tag{5.16}$$

If we condition on the first observation, then the likelihood ratio test statistics are related by the identity: $LR_{CC} = LR_{UC} + LR_{IND}$

Christoffersen's basic framework is limited in that it only deals with first-order dependence in the $\{I_t\}$ series. It would fail to reject a $\{I_t\}$ series that does not have first-order Markov dependence but does exhibit some other kind of dependence structure (e.g. higher order Markov dependence or periodic dependence). Three recent papers, Christoffersen and Diebold (2000), Clements and Taylor (2003) and Engle and Manganelli (2004), generalize this observation. These papers suggest that a regression of the I_t series on its own lagged values and some other variables of interest, such as day-dummies or the lagged observed returns (the periodic component is *S*), can be used to test for the existence of various forms of dependence structure that may be present in the $\{I_t\}$ series. Under this framework, the conditional efficiency of the I_t process can be tested by testing the joint hypothesis:

$$H : \Phi = 0, \alpha_0 = p. \tag{5.17}$$

where

$$\Phi = [\alpha_1, ...\alpha_S, \mu_1, ..., \mu_S]',$$

in the regression

$$I_t = \alpha_0 + \sum_{s=1}^{S} \alpha_s I_{t-s} + \sum_{s=1}^{S} \mu_s D_{s,t} + \varepsilon_t . \tag{5.18}$$

where $t = S+1, S+2, ..., T$ and $D_{s,t}$ are explanatory variables.

The hypothesis (5.17) can be tested by using an F-statistic in the usual OLS framework.

We then employ the regression-based tests of Clements and Taylor (2003) to test for both independence and correct conditional coverage. This is done by performing an OLS regression of the I_t series on its lagged value. For the independence test, we set the following hypothesis:

$$H : \Phi = 0. \tag{5.19}$$

Similarly, we test for correct conditional coverage by testing the joint hypothesis in (5.17). Both tests are again conducted for 95% and 99% VaR coverage for all out-of-sample VaR measures constructed for the 100-day out-of-sample forecasting period.

5.3 Results

This section reports the findings of the estimations and tests performed on the various available GARCH-based models as well as the VaR models described above. The discussion of results will be conducted in the following stages. In the first stage, the discussion will focus on the evaluation of the model fits of the periodic GARCH and non-periodic GARCH models as well as the naïve model obtained from the initial sample estimates.

The objective of this exercise is to ascertain whether periodic GARCH models produce superior model fits to the non-periodic GARCH models and the naïve model. A comparison of ranking with the results obtained in Chapter 4 is then carried out with regard to which volatility modelling approach produces the best model fit. At this stage, we want to see whether there is any consistency in the performance ranking for the thirteen volatility modelling approaches for both the 406-day (performed in Chapter 4) and the 306-day (performed in this chapter) in-sample estimates.

The second stage of the analysis reports the forecasting performance produced by each of the GARCH-based models and the naive model, using the MSFE and the MAFE error statistics. Of major interest from this section is to ascertain whether the models with better fits (evaluated over the in-sample period) produce better forecasting performance than the less well fitted models. The relationship between model fit in the initial estimation periods and the ensuing forecasting performances obtained would be valuable for any future work in the field. If it is determined that better model fit translates to superior forecasting performances, then what one should do is to evaluate all available models extensively in the initial estimate periods and choose the best performing model. The best model should then be able to produce the best volatility forecasts.

The third stage considers the quality of the various forecasts available using the DM asymptotic test, followed by the HLN forecast encompassing test. We compare the quality of forecasts produced by each of the GARCH-based models with the forecast generated by the naive model. Next we compare the forecasts produced by each of the volatility modelling approaches. The aim is not only to determine whether the PGARCH forecasts are superior in quality compared to those produced by non-periodic GARCH models; we also wish to ascertain whether there is any statistically significant difference in terms of quality among the available PGARCH-based models forecasts.

In the final stage, we discuss the performance of the VaR measures generated by the various GARCH-based models. This covers the three tests under the framework developed by Christoffersen (1998), as well as the results of the regression-based tests proposed by Clements and Taylor (2003).

5.3.1 Model Fit

The results of the in-sample estimates for the thirteen competing volatility modelling approaches are presented in Table 5.2. The best model fit is produced by approach T10, followed by those produced by approaches T2, T12, T11, T13, T5, T3, T6, T8, T4, T7, T9 and T1, respectively. It is clear that all the volatility modelling approaches that utilize PGARCH-based models exhibit superior model fit compared to the non-periodic GARCH model. The results are consistent with the results obtained in Chapter 4 (please refer to Table 4.16). However, when we compare the ranking of volatility modelling approaches, we find that with the exception of the T9 and T11 approaches, there is a change in the order of ranking for all approaches. The ranking of best performing approaches from Chapter 4 is as follows: T2, T12, T10, T4, T11, T13, T5, T3, T6, T7, T8, T9 and T1, respectively. It is clear now that approach T10, which is the jointly estimated spline version of the PGARCH model, is the best performing approach in the in-sample estimates. It is again encouraging to see that both the business time based spline version and the calendar time based spline version of the PGARCH models (both jointly estimated and two-step estimation) show strong performance in the 306-day in-sample period. The ranking of the FFF version of the PGARCH models in this study has also changed. Now, we find that the jointly estimated approaches (T6 and T8) are superior in terms of model fit to the two-step approaches (T7 and T9) where previously, in the 406-day in-sample estimates the results indicate that the T7 approach is superior to the T8 approach. There is evidence that the choice of in-sample estimation period does influence the ranking of performance of the volatility modelling approaches.

5.3.2 Forecast Performance

Regarding forecasting performance, the results presented in Table 5.3 show that both PGARCH and non-periodic GARCH models perform much better than the naive model. It is clear that all available GARCH models provide smaller forecast error statistics (both for MSFE and MAFE) than those for the naive model. Comparison of forecasting performance among the thirteen volatility modelling approaches, however, has to be evaluated separately according to the choice of forecast error applied. If the MSFE statistic is applied, the ranking shows that approach T10 (the jointly estimated spline version of the PGARCH model) exhibits the best forecasting performance. The subsequent best performances are shown by approaches T12, T2, T11, T5, T13, T6, T8, T3, T7, T9, T4 and T1, respectively. If we take the MAFE as the choice of forecast error statistics, we find that approach T12 now produces the best forecasting performance. This is followed by the performances of approaches T10, T2, T6, T3, T4, T11, T8, T13, T5, T9, T7 and T1, respectively. The ranking order now is completely different from that obtained when the MSFE is applied. It is clear, however, that the first three positions have been occupied by the same approaches under both MSFE and MAFE statistics. It is also demonstrated that the jointly estimated spline version of the PGARCH model (both calendar time and business time based) produces better results than the jointly estimated full dummy version of the PGARCH model, even though the T12 approach produced inferior model fit to the T2 approach in the in-sample estimation period. Regarding the best GARCH model to use in each approach, the results indicate that at least for the Malaysian market, the EGARCH model is the most appropriate GARCH-based model to be considered in forecasting volatility, since the EGARCH model is used by the three top volatility modelling approaches here. There is also a strong indication that using the spline variables in volatility fitting and forecasting is more appropriate and therefore is recommended for the Malaysian market.

It is clear from the above results that there is a departure in consistency when we compare the rankings of in-sample estimation performances with the forecasting performances of the thirteen volatility modelling approaches considered in this study. The results suggest that the approach that produces the best in-sample model fit does not necessarily produce the best forecasting performance. However, there are indications that volatility modelling approaches that exhibit superior in-sample model fit statistics generally produce better forecasting performances than the approaches which show poor in-sample model fit statistics. The results also demonstrate that on the whole, the PGARCH-based models produce superior forecasting performances to the non-periodic GARCH models. This is clearly evident when both the MSFE and the MAFE are applied. Therefore, the incorporation of periodic components in the conditional volatility process is strongly recommended because it does improve the forecasting performance of the volatility model.

The plots of the average of 5-minute out-of-sample forecasts of individual volatility modelling approaches against the realized volatility are presented in Figures 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9, 5.10, 5.11, 5.12 and 5.13 respectively. Figure 5.14 presents the plots of all forecasts for the volatility modelling approaches T1-T13 against the realized volatility. It is clear that the T2, T10 and T12 approaches, which use the EGARCH model, exhibit the best fitted volatility as well as doing an adequate job in modelling the double U-shaped intraday periodicity, which is observable through the plot of the *ex post* realized volatility.

5.3.3 Forecast Quality

The results are presented in Tables 5.4 to 5.6. The DM tests performed using both the MSFE and the MAFE as the loss functions produce significant results for all comparisons of the naive model and each of the volatility modelling approaches, T1 to T13. The test statistics obtained are significant at the 1% level. The results suggest that the quality of the forecasts produced by all the volatility modelling approaches is superior to those generated by the naive model. Similar results are also obtained when we assess the quality of forecasts for all comparisons of the non-periodic GARCH approach (T1) and each of the PGARCH-based modelling approaches (T2 to T13).

The results are not as clear-cut as when the same tests are performed among competing pairs of the PGARCH-based models forecasts. If we consider the MSFE loss function, we find that the following pairs provide insignificant results: T2 v. T5, T2 v. T11, T2, v. T12, T3 v. T4, T3 v. T6, T3 v. T7, T3 v. T8, T3 v. T9, T4 v. T9, T5 v. T6, T5 v. T8, T5 v. T9, T5 v. T11, T5 v. T13, T6 v. T8, T6 v. T13, T7 v. T8, T7 v. T9, T8 v. T9, T8 v. T13, T10 v. T12, T11 v. T12 and T11 v. T13. Similarly, if we consider the MAFE loss function, the following pairs produce insignificant results: T2 v. T10, T2 v. T12, T3 v. T4, T3 v. T6, T5 v. T8, T5 v. T9, T5 v. T13, T7 v. T9, T8 v. T11, T8 v. T13, T10 v. T12 and T11 v. T13. All forecast comparisons are statistically insignificant at the 5% level, and therefore suggest that the quality of forecasts among these competing PGARCH-based models is similar.

It is also clear that more insignificant results are obtained when we use the MSFE rather than the MAFE as the loss function. A possible explanation for these findings could be attributed to the components of the metric used in the tests, which is the MSFE loss function. The MSFE loss function employs squared forecast errors in its computation and this complicates the task of forecast evaluation when used in the Diebold and Mariano (1995) test, because the square of a variance error is the 4th power of the same error measured from standard deviation. Poon and Granger (2003) highlight this complication when forecast errors are measured from variances. They argue that the confidence interval of the mean error statistic can get very wide and the situation can be made worse if the variance errors are squared. This leads to extremely noisy "quality" estimates that in turn make the small differences between squared forecast errors of models tested indistinguishable from each other.

The overall results indicate that regardless of whether we use the MSFE or the MAFE to evaluate the quality of competing forecasts, the quality of forecasts of the

PGARCH-based approaches is superior to the quality of forecasts of both the naive model and the non-periodic GARCH-based model. This suggests that the inclusion of periodic components in the forecasting exercise does provide benefits in terms of superior model fits and forecast quality.

The results of the HLN forecast encompassing tests are presented in Tables 5.6 to 5.7. The overall results seemed to confirm the above findings. The HLN test statistics are significant at 1% significant level for all comparisons of forecasts between the naive model and the individual volatility modelling approaches T1 to T13. Similar findings are observed when we compare the quality of forecasts of the PGARCH models against the non-periodic GARCH models. Therefore, we conclude that in this instance, for all forecast quality comparisons under the encompassing test, we reject the null hypothesis that the forecasts from competing models encompass each other, i.e., the competing forecasts embody no useful information that is missing in the preferred forecasts. The results suggest that there is statistically a difference in the quality of forecasts generated by the naive, the non-periodic GARCH and the PGARCH models. This is in line with the findings of the DM tests above.

When we look at the comparisons of forecasts among the PGARCH-based modelling approaches, we find the following competing pairs produce results that are insignificant at the 5% level: T2 v. T3, T2 v. T4, T2 v. T5, T2 v. T6, T2 v. T7, T2 v. T8, T2 v. T9, T2 v. T11, T2 v. T13, T5 v. T6, T5 v. T8, T6 v. T7, T6 v. T8, T6 v. T9, T7 v. T8, T7 v. T9, T10 v. T11, T10 v. T12, T10 v. T13 and T11 v. T13. Therefore, we conclude that the quality of forecasts for these twenty comparisons is similar.

5.3.4 Assessment of VaR Performance

As mentioned earlier, the GARCH models are widely used for the management of risk. It is thus important to assess their quality through the various VaR measures discussed below. We begin with the assumption of conditional normal distribution in the return series and the results of the coverage tests for the thirteen modelling approaches and the *RiskMetrics* model. We then report the results of the coverage tests for the Student t distribution assumption, with 4 and 24 degrees of freedom.

5.3.4.1 Test for "Correct Unconditional Coverage" H_0: $\hat{f} = \alpha$

99% VaR Forecast

The main criterion is to achieve a probability of failure $\hat{f}$ equal to the desired level, i.e. α =0.01. The results of backtesting the VaR models for the 100-day out-of-sample forecasting period assuming normal distribution are shown in Table 5.8. The Christoffersen (1998) test rejects all the volatility modelling approaches except T1. The likelihood ratio statistics for approaches T2 to T13, and the *RiskMetrics*, are all highly significant at the 1% level. This suggests that approaches T2-T13, and the *RiskMetrics,* do not have the correct unconditional coverage property. However, the likelihood ratio statistics for T1 is not statistically significant at the 1% level of significance, and therefore appears to satisfy the required coverage.

We next look at the results for the Student t distribution with 4 degrees of freedom assumption, which are shown in Table 5.14. It is clear that all results are statistically significant at the 1% level. Therefore, the results suggest that all modelling approaches (inclusive of the *RiskMetrics*), fail the correct unconditional coverage tests. However, more favourable results are obtained when we consider the results of the Student t distribution with 24 degrees of freedom assumption, which are reported in Table 5.20. With the exception of the T3, T4, T5, T13 and the *RiskMetrics* approaches, the results of the coverage test for the other nine modelling approaches are not significant at the 1% level. Therefore, it appears that the nine modelling approaches satisfy the correct unconditional coverage property.

95% VaR Forecast

We turn to the results for the normal distribution assumption, which are presented in Table 5.9. The likelihood ratio statistics for the T4, T8, T11, T12 and T13 volatility modelling approaches are statistically insignificant at the 5% level. Therefore, it appears that only these five approaches have the correct unconditional coverage. The rest, including the *RiskMetrics* approach, fail the correct unconditional test.

Similarly, the results of the coverage tests for the Student t distribution with assumptions of 4 degrees of freedom and 24 degrees of freedom are presented in Table 5.15 and Table 5.21, respectively. All results with the exceptions of the T4 and T5 approaches (for the 24 degrees of freedom) are statistically significant at the 5% level. All the modelling approaches fail the coverage test, and therefore, none of the modelling approaches appears to have the necessary coverage.

5.3.4.2 Test for "Independence"

First, we focus on the results for the normal distribution assumption for the 99% and 95% VaR models, which are presented in Table 5.10 and Table 5.11, respectively. For the 99% VaR coverage, we find that all approaches T1-T13 and the *RiskMetrics* fail the independence test. The F-statistics obtained for these approaches are highly significant at 1% level. Therefore these approaches do not appear to have the required coverage property. For the 95% VaR coverage, only the T1, T2 and the T10 approaches pass the independence test. The F-statistics of all other approaches are significant at the 5% level. Therefore, we conclude that only the T1, T2 and T10 approaches satisfy the independence coverage criteria.

Next, we turn to the results of the 99% and 95% VaR independence tests for the Student t distribution, assuming 4 degrees of freedom, which are presented in Table 5.16 and Table 5.17, respectively. For the 99% VaR coverage, we find that with the exception of the T4, T6 and the *RiskMetrics* approaches, all the modelling approaches pass the independence test. The F-statistics for these approaches are not statistically significant at the 1% level, and therefore, they appear to have the required coverage property. For the 95% VaR coverage, we find that with the exceptions of the T4, T7, T9 and T11 approaches, all the other approaches including the *RiskMetrics* approach appear to have the necessary coverage. The F-statistics for these approaches are not significant at the 5% level.

Next, we look at the results of the 99% and 95% VaR independence tests for the Student t distribution, assuming 24 degrees of freedom, which are reported in Table 5.22 and Table 5.23, respectively. For the 99% VaR coverage, we find that only approach T6 passes the independence tests. The F-statistics is insignificant at the 1% level. The rest of the modelling approaches do not appear to have the independence property. For the 95% VaR coverage, we find that only the T1, T2, T7, T10 and the T12 approaches have the independence property. The results for the rest of the modelling approaches are statistically significant at the 5% level, and therefore, they fail the independence test.

5.3.4.3 Test for "Correct Conditional Coverage"

We analyse the results of both the 99% and 95% VaR measures for the normal distribution assumption, which are presented in Table 5.12 and Table 5.13, respectively. For the 99% VaR coverage, we find that the regression-based tests reject the adequacy of all models conclusively, with all F-statistics obtained showing significance at the 1% level. The results exhibit the existence of significant lagged dependence in the failure process. Given these results, we conclude that none of the VaR models are appropriate for the KLCI returns. The results are similar when we focus on the 95% VaR coverage. We find that all approaches produce results that are statistically significant at the 5% level. Therefore, these approaches do not have the correct conditional property.

Next, we turn to the results of the 99% and 95% coverage for the Student t distribution with 4 degrees of freedom. The results are reported in Tables 5.18 and 5.19 respectively. The results obtained for both the 99% and 95% VaR coverage indicate that all modelling approaches fail the correct conditional test. The results are statistically significant at the 1% level. Therefore, none of the modelling approaches has the required VaR coverage.

Finally, we analyse the results of the 99% and 95% coverage for Student t distribution with 24 degrees of freedom, which are reported in Tables 5.24 and 5.25 respectively. For the 99% VaR coverage, we find that all approaches do not have the required correct conditional coverage property. The results for all approaches are significant at the 1% level. The results for the 95% VaR coverage indicate that all the approaches again fail the correct conditional coverage test. All results are highly statistically significant at the 1% level. Therefore, none of the modelling approaches has the correct conditional coverage property.

5.4 Conclusion

In this chapter, we have evaluated the forecasting performance of the thirteen volatility approaches that utilize the PGARCH-based models and the non-periodic GARCH specification described in Chapter 4. We find that approaches that use PGARCH-based models achieve better model fit compared to the non-periodic GARCH models when in-sample analysis is performed. This is evinced by the superior AIC and SIC statistics produced in the in-sample initial period generated by the PGARCH-based models. Due to the superior model fit, the PGARCH-based models (used by approaches T2 to T13) also produce better forecast performance than the non-periodic GARCH model used in approach T1. Specifically, the MSFE and MAFE values show that PGARCH forecasts produce smaller forecast errors than the forecast error associated with the non-periodic GARCH model. The results of both forecast error statistics also indicate that all GARCH-based models utilized by the thirteen volatility modelling approaches produce smaller forecast errors than those obtained for the naive model. The results suggest that the inclusion of periodic components when modelling volatility would improve the forecasting performance of the particular volatility model. However, as shown in section 5.3.1, the model or the approach that produces the best in-sample statistics does not necessarily produce the best forecasting performance. The overall results, however, demonstrate that the approaches that make use of the spline variables in the conditional volatility process tend to perform better in a forecasting exercise. More specifically, the best out-of-sample forecasting performance incidentally is produced by the jointly estimated spline version of the PGARCH model (T10 – based on MSFE) and the jointly estimated augmented spline version of the PGARCH model (T12 – based on MAFE). This is not surprising considering that these two approaches are consistently highly ranked in the in-sample estimates and in the ranking of the best performing modelling approaches in Chapter 4.

With regard to the asymptotic tests of Diebold and Mariano (1995) for "equal accuracy" and the encompassing forecast tests of Harvey, Leybourne and Newbold (1998), we find strong evidence that the quality of forecasts produced by both the PGARCH-based models and non-periodic GARCH models are superior to the naive model, which is based on historical variance. Furthermore, we also find evidence that the forecasts produced by the PGARCH-based models are superior to the forecasts produced by the non-periodic GARCH models. The results from comparing the alternative forecasts among the PGARCH-based models also indicate that the quality of forecasts of the top three best forecast performing models are the same. Similarly, there is a difference in terms of quality when the forecasts of these top three performers are compared with the forecasts of the less well performing PGARCH-based models. The strong performance of the PGARCH-based models over the non-periodic GARCH models and the naive model is an encouraging result, in view of the statistical significance provided by the two asymptotic tests in relation to the evaluation of forecasts. The results also reinforce the virtue of the PGARCH structure, which has now been proven not only capable of providing superior model fit (as evinced in Chapter 4), but more importantly, of generating genuinely superior forecasts, which is the real test of a good forecasting model. However, it is important to appreciate that the two asymptotic tests (the DM and HLN tests) do not produce similar results. Even within the DM tests, we have seen that different results can be obtained in evaluating the same pair of competing alternative forecasts. The results obtained in this chapter, for example, indicate that the statistical significance of the results very much depends on the type of loss function used. This is evinced in some of the different results obtained when we apply both the MAFE and MAFE loss functions on the same pair of competing alternative forecasts. Part of the problem could be the effect of squaring the squared forecast errors when the MSFE is used as the loss function. Therefore, as far as the DM test is concerned, we believe that the better choice is to use the MAFE as the loss function.

We now turn to the performance of VaR models constructed based on the forecasts of the thirteen volatility modelling approaches and the *RiskMetrics* model. We have strong reservations about getting favourable results when the normality assumption is applied due to the fat-tailed distribution of the KLCI returns series used in this study. In particular, the *RiskMetrics* approach makes the very strong assumption that returns are conditionally[9] normally distributed. Therefore, we provide alternative distribution assumptions in the form of the Student t distribution, initially with 4 degrees of freedom to control for a very fat-tailed distribution, and later with 24 degrees of freedom to account for a less severe effect of the fat tails. The results for the normal distribution indicate that for the 99% VaR coverage, the effect of fat tails is very strong, and as expected, the VaR is seriously underestimated. This is perhaps the reason for the poor results obtained for all modelling approaches in the correct conditional coverage test. Similar results are obtained when we consider the 95% VaR models. We find none of the approaches pass the independence and the correct conditional coverage tests, and are therefore these approaches do not have the necessary coverage property.

When we analyse the overall results for the 99% and 95% VaR models using a Student t distribution with 4 degrees of freedom, the results indicate that this distribution is not suitable, due to its failure to accommodate the fat-tail effects. The

[9] "Conditional" here means conditional on the information set at time t, which usually consists of the past return series available at time t.

use of a Student t distribution with 24 degrees of freedom does not appear to mitigate the effect of fat tails, at both the VaR coverage. We find that at the 99% VaR coverage, all the PGARCH-based models and the *RiskMetrics* model fail both the independence and the correct conditional coverage requirements, and therefore do not produce accurate VaR models. The same could be said for the results for the 95% VaR coverage, where we find again that all the modelling approaches fail the correct conditional coverage test. Therefore we conclude that none of the VaR models constructed in this study are economically reliable.

The work done in this chapter completes the second out of the three investigations towards a better understanding of the dynamics of intraday volatility on the KLSE. In this chapter, we have focused on the measurement and evaluation of the performance and quality of various volatility forecasts produced by competing modelling approaches using high frequency data. More specifically, we have shown the superiority of the PGARCH-based models in producing superior forecasts. In the next chapter, we will continue modelling and forecasting using high frequency data. We will investigate a new volatility measure in the form of integrated realized volatility, which can be constructed by summing the intraday high-frequency squared returns from different return sampling frequencies. Similar to the works carried out in this chapter, we will evaluate the performance and the quality of forecasts, as well as the accuracy of the VaR models generated by the ARMA and the GARCH models.

Table 5.2: Comparison of Volatility Modelling Approaches – KLCI
306 In-sample Estimation Period

This table describes the in-sample parameter estimates of the five volatility modelling approaches described and denoted in the text as approaches T1, T2, T3, T4, and T5. The numbers in parentheses are Bollerslev-Wooldridge robust standard errors. The log likelihood (LL), the Akaike Information Criterion (AIC), and the Schwarz Information Criterion (SIC) are also given. The significance of these estimates is denoted by ** (1% significance) and * (5% significance).

	T1 EGARCH	T2 EGARCH	T3 GARCH	T4 EGARCH	T5 GARCH
Mean Equation					
C	-0.0003	-0.0006	0.0014	-0.0002	-0.0004
	(0.0006)	(0.0005)	(0.0062)	(0.0007)	(0.0060)
Volatility Equation					
α_1			0.0301**		0.0349**
			(0.0030)		(0.0034)
β_1			0.9641**		0.9586**
			(0.0035)		(0.0039)
|RES|/SQR[GARCH](1)	0.2749**	0.0859**		0.0884**	
	(0.0202)	(0.0098)		(0.0113)	
RES/SQR[GARCH](1)	-0.0109	8.63E-05		-0.0027	
	(0.0102)	(0.0045)		(0.0059)	
EGARCH(1)	0.9383**	0.9928**		0.9929**	
	(0.0082)	(0.0017)		(0.0021)	
C	-0.4848**	0.2191**	0.0056**	-0.0709**	0.0062**
	(0.0417)	(0.0238)	(0.0018)	(0.0140)	(0.0019)
Model Fit					
LL	22140.12	23330.17	23041.034	22584.34	23084.81
AIC	-1.9551	-2.0592	-2.0347	-1.9939	-2.0386
SIC	-1.9533	-2.0535	-2.0333	-1.9907	-2.0372
RANKING	**13**	**2**	**7**	**10**	**6**

**Table 5.2: Comparison of Volatility Modelling Approaches – KLCI
306 In-sample Estimation Period (continued)**

This table describes the in-sample parameter estimates of the six volatility modelling approaches described and denoted in the text as approaches T6, T7, T8, T9, T10 and T11. The numbers in parentheses are Bollerslev-Wooldridge robust standard errors. The log likelihood (LL), the Akaike Information Criterion (AIC), and the Schwarz Information Criterion (SIC) are also given. The significance of these estimates is denoted by ** (1% significance) and * (5% significance).

	T6 EGARCH	T7 TGARCH	T8 EGARCH	T9 TGARCH	T10 EGARCH	T11 TGARCH
Mean Equation						
C	-8.59E-05	0.0005	-8.59E-05	-0.0002	-0.0004	-0.0029
	(0.0006)	(0.0058)	(0.0005)	(0.0057)	(0.0005)	(0.0059)
Volatility Equation						
α_1		0.0372**		0.0412**		0.03173**
		(0.0044)		(0.0044)		(0.0038)
β_1		0.9544**		0.9487**		0.9633**
		(0.0043)		(0.0045)		(0.0032)
γ		0.0012		0.0029		-0.0009
		(0.0053)		(0.0055)		(0.0043)
|RES|/SQR[GARCH](1)	0.0895**		0.0897**		0.0871**	
	(0.0114)		(0.0113)		(0.0017)	
RES/SQR[GARCH](1)	-0.0016		-0.0016		3.50E-05**	
	(0.0055)		(0.0058)		(0.0012)	
EGARCH(1)	0.9914**		0.9934**		0.9921**	
	(0.0021)		(0.0026)		(0.0004)	
C	-0.1045**	0.0082**	-0.1085**	0.0094**	-0.2952**	0.0053**
	(0.0149)	(0.0019)	(0.0156)	(0.0018)	(0.0056)	(0.0015)
Model Fit						
LL	22824.15	22553.66	22822.13	22469.87	23359.67	23279.27
AIC	-2.0156	-1.9916	-2.0149	-1.9842	-2.0617	-2.0557
SIC	-2.0124	-1.9898	-2.0117	-1.9824	-2.0557	-2.0539
RANKING	**8**	**11**	**9**	**12**	**1**	**4**

Table 5.2: Comparison of Volatility Modelling Approaches – KLCI
306 In-sample Estimation Period (continued)

This table describes the in-sample parameter estimates of the six volatility modelling approaches described and denoted in the text as approaches T6, T7, T8, T9, T10 and T11. The numbers in parentheses are Bollerslev-Wooldridge robust standard errors. The log likelihood (LL), the Akaike Information Criterion (AIC), and the Schwarz Information Criterion (SIC) are also given. The significance of these estimates is denoted by ** (1% significance) and * (5% significance).

	T12 EGARCH	T13 TGARCH
Mean Equation		
C	-0.0007	-0.0037
	(0.0005)	(0.0058)
Volatility Equation		
α_1		0.0311**
		(0.0037)
β_1		0.9649**
		(0.0033)
γ		-0.0017
		0.0042
|RES|/SQR[GARCH](1)	0.0867**	
	(0.0017)	
RES/SQR[GARCH](1)	0.0002	
	(0.0012)	
EGARCH(1)	0.9927**	
	(0.0004)	
C	-0.3479**	0.00478*
	(0.0229)	(0.0013)
Model Fit		
LL	23299.80	23265.57
AIC	-2.0565	-2.0545
SIC	-2.0539	-2.0527
RANKING	**3**	**5**

Table 5.3: Forecast Performance 100-Day Out-of-sample Forecasting Period-PGARCH

The following table reports two forecast error statistics for the forecasts produced by the various volatility modelling approaches listed below. The results are based on a one-step-ahead forecast covering a 100-day out-of-sample period. The errors computed are the mean squared forecast error (MSFE) and the mean absolute forecast error (MAFE).

		MSFE	Ranking based on MSFE	MAFE	Ranking based on MAFE
	Naive Model	0.000412		0.014005	
	Volatility Modelling Approach:				
T1	Non-periodic GARCH model	0.000290	13	0.006158	13
T2	Jointly-estimated full dummy version of the PGARCH model	0.000265	3	0.005297	3
T3	Two-step full dummy version of the PGARCH model	0.000281	9	0.005547	5
T4	Jointly-estimated partial dummy version of the PGARCH model	0.000293	12	0.005584	6
T5	Two-step partial dummy version of the PGARCH model	0.000274	5	0.005705	10
T6	Jointly estimated FFF version of the PGARCH model	0.000277	7	0.005468	4
T7	Two-step FFF version of the PGARCH model	0.000283	10	0.005951	12
T8	Jointly estimated Augmented FFF version of the PGARCH model	0.000279	8	0.005674	8
T9	Two-step Augmented FFF version of the PGARCH model	0.000285	11	0.005879	11
T10	Jointly estimated Spline version of the PGARCH model	0.000261	1	0.005287	2
T11	Two-step Spline version of the PGARCH model	0.000267	4	0.005690	7
T12	Jointly estimated Augmented Spline version of the PGARCH model	0.000264	2	0.005282	1
T13	Two-step Augmented Spline version of the PGARCH model	0.000276	6	0.005701	9

Table 5.4: Comparing Forecast quality - 100-Day Out-of-sample Forecasting Period

This table reports the results of the Diebold and Mariano (1995) asymptotic test for forecast quality evaluation. The null hypothesis is that the forecasts generated based on unconditional variance (M1) are of the same quality as the forecasts generated by the various volatility modelling approaches listed in Table 5.4 above. The alternative hypothesis adopted is that the forecasts produced by the volatility modelling approaches are superior to the forecasts of M1. The results are based on a one-step-ahead forecast covering a 100-day out-of-sample period. The test is implemented with the mean squared forecast error (MSFE) and the mean absolute forecast error (MAFE) as the loss functions. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

	Metric	
	MSFE	MAFE
Naive Model (M1)		
Naive Model v. Volatility Approach		
M1 v. T1	28.9677**	111.5225**
M1 v. T2	33.9285**	122.3850**
M1 v. T3	13.2278**	88.5388**
M1 v. T4	27.3170**	119.1960**
M1 v. T5	14.6689**	92.8531**
M1 v. T6	54.2777**	126.7412**
M1 v. T7	24.4348**	113.6956**
M1 v. T8	55.6680**	118.3783**
M1 v. T9	22.7638**	114.3227**
M1 v. T10	27.2498**	120.7394**
M1 v. T11	18.9399**	106.7386**
M1 v. T12	32.2998**	123.4641**
M1 v. T13	14.1643**	89.9590**

Table 5.5: Comparing Forecast Quality between Volatility Modelling Approaches

This table reports the results of the Diebold and Mariano (1995) asymptotic test for forecast quality evaluation. The null hypothesis is that the forecast generated based on a volatility modelling approach is of the same quality as the forecast generated by another competing volatility modelling approach. The alternative hypothesis adopted is that the forecast produced by a volatility modelling approach is superior to the forecast generated by a competing volatility modelling approach. The results are based on a one-step-ahead forecast covering a 100-day out-of-sample period and comparisons of forecast quality are performed among the eleven modelling approaches described in Table 5.4. The tests are implemented with the mean squared forecast error (MSFE) and the mean absolute forecast error (MAFE) as the loss functions. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

Comparison of Modelling Approaches	MSFE	MAFE
T1 v. T2	4.3308**	19.4479**
T1 v. T3	2.8766**	10.5286**
T1 v. T4	2.7168**	13.4282**
T1 v. T5	2.5245**	7.1157**
T1 v. T6	4.0038**	20.5943**
T1 v. T7	3.7369**	19.1887**
T1 v. T8	3.2555**	13.0602**
T1 v. T9	3.6628**	22.1336**
T1 v. T10	4.1741**	19.3239**
T1 v. T11	3.2733**	15.1879**
T1 v. T12	4.3373**	20.0497**
T1 v. T13	3.1635**	3.9132**
T2 v. T3	-2.5590**	-5.2430**
T2 v. T4	-3.9073**	-6.2292**
T2 v. T5	-1.5812	-9.9015**
T2 v. T6	-2.9737**	-6.3934**
T2 v. T7	-2.9544**	-6.4356**
T2 v. T8	-3.8773**	12.0068**
T2 v. T9	-2.6002**	-4.5452**
T2 v. T10	1.5869	0.9565
T2 v. T11	-0.4185	-4.3717**
T2 v. T12	1.4933	1.4766
T2 v. T13	-2.1822*	-14.0415**
T3 v. T4	-0.9478	-0.5436
T3 v. T5	2.0396*	-5.4331**
T3 v. T6	0.3821	1.4045
T3 v. T7	-0.3745	10.5847**
T3 v. T8	0.2174	2.2318*
T3 v. T9	-0.3130	-10.6115**
T3 v. T10	3.5683**	5.5488**
T3 v. T11	2.6448**	2.0499*
T3 v. T12	2.7939**	5.5148**
T3 v. T13	1.9620*	-11.0735**

Table 5.5: Comparing Forecast Quality between Volatility Modelling Approaches (continued)

This table reports the results of the Diebold and Mariano (1995) asymptotic test for forecast quality evaluation. The null hypothesis is that the forecast generated based on a volatility modelling approach is of the same quality as the forecast generated by another competing volatility modelling approach. The alternative hypothesis adopted is that the forecast produced by a volatility modelling approach is superior to the forecast generated by a competing volatility modelling approach. The results are based on a one-step-ahead forecast covering a 100-day out-of-sample period and comparisons of forecast quality are performed among the eleven modelling approaches described in Table 5.4. The tests are implemented with the mean squared forecast error (MSFE) and the mean absolute forecast error (MAFE) as the loss functions. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

Comparison of Modelling Approaches	MSFE	MAFE
T4 v. T5	4.5719**	-3.8299**
T4 v. T6	4.9923**	3.2349**
T4 v. T7	2.2520*	-8.4924**
T4 v. T8	3.6295**	-2.0168*
T4 v. T9	1.2053	-2.9685**
T4 v. T10	3.7249**	6.2097**
T4 v. T11	2.7515**	-2.3126*
T4 v. T12	3.9059**	6.6777**
T4 v. T13	3.2583**	-4.6407**
T5 v. T6	-0.3819	4.4843**
T5 v. T7	-3.4126**	-2.4923**
T5 v. T8	-0.5989	0.5751
T5 v. T9	-3.4440	-1.6221
T5 v. T10	2.5618**	10.3286**
T5 v. T11	0.8830	4.8839**
T5 v. T12	2.8022**	10.1695**
T5 v. T13	-1.9242	1.8886
T6 v. T7	-1.9709*	-13.5379**
T6 v. T8	-0.1781	-2.5288**
T6 v. T9	-2.1809*	-8.6218**
T6 v. T10	2.9136**	6.0267**
T6 v. T11	2.4565**	-2.5656**
T6 v. T12	3.0074**	7.2359**
T6 v. T13	0.01278	-8.0263**

Table 5.5: Comparing Forecast Quality between Volatility Modelling Approaches (continued)

This table reports the results of the Diebold and Mariano (1995) asymptotic test for forecast quality evaluation. The null hypothesis is that the forecast generated based on a volatility modelling approach is of the same quality as the forecast generated by another competing volatility modelling approach. The alternative hypothesis adopted is that the forecast produced by a volatility modelling approach is superior to the forecast generated by a competing volatility modelling approach. The results are based on a one-step-ahead forecast covering a 100-day out-of-sample period and comparisons of forecast quality are performed among the eleven modelling approaches described in Table 5.4. The tests are implemented with the mean squared forecast error (MSFE) and the mean absolute forecast error (MAFE) as the loss functions. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

Comparison of Modelling Approaches	MSFE	MAFE
T7 v. T8	0.2496	3.5811**
T7 v. T9	-1.0245	1.0790
T7 v. T10	2.7936**	6.4912**
T7 v. T11	2.6474**	3.2487**
T7 v. T12	2.8790**	6.8542**
T7 v. T13	2.6654**	5.5613**
T8 v. T9	0.1543	-5.6248**
T8 v. T10	3.5796**	11.2892**
T8 v. T11	2.6685**	1.9180
T8 v. T12	3.8493**	6.0678**
T8 v. T13	0.1787	-1.1813
T9 v. T10	3.6175**	4.6494**
T9 v. T11	2.5878**	2.4790**
T9 v. T12	3.4181**	4.9648**
T9 v. T13	1.9782*	2.6522**
T10 v. T11	-2.2991*	-4.9265**
T10 v. T12	-1.1547	0.4329
T10 v. T13	-3.2016**	-14.1042**
T11 v. T12	0.6014	4.6423**
T11 v. T13	-1.3237	-1.1903
T12 v. T13	-2.4173**	-14.4052**

Table 5.6: Comparing Forecast Quality – Forecast Encompassing Test

This table reports the results of the asymptotic "forecast encompassing" test of Harvey et al. (1998) using a 5% significance level. The test statistic is denoted by HLN. The null hypothesis is that if forecasts from model 1 encompass the forecasts from model 2, then the covariance between e_{1t} and e_{1t}-e_{2t} will be negative or zero (e_{1t} and e_{2t} are the two sets of forecast errors from model 1 and model 2 respectively). The alternative hypothesis is that the forecasts from model 1 do not encompass those from model 2, in which case the covariance between e_{1t} and e_{1t}-e_{2t} will be positive. The results are based on a one-step-ahead forecast covering a 100-day out-of-sample period. Forecast quality comparisons are made with unconditional variance (M1) with the various volatility modelling approaches listed in Table 5.4. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

	Metric
	HLN
Naive Model (M1)	
Naive Model v. Volatility Approach	
M1 v. T1	192.5334**
M1 v. T2	224.8631**
M1 v. T3	129.0693**
M1 v. T4	336.6812**
M1 v. T5	131.8517**
M1 v. T6	272.5332**
M1 v. T7	199.5606**
M1 v. T8	212.3302**
M1 v. T9	206.0263**
M1 v. T10	212.8964**
M1 v. T11	165.8173**
M1 v. T12	229.4463**
M1 v. T13	124.0719**

Table 5.7: Comparing Forecast Quality between Volatility Modelling Approaches – Forecast Encompassing Test

This table reports the results of the asymptotic "forecast encompassing" test of Harvey et al. (1998) using a 5% significance level. The test statistic is denoted by HLN. The null hypothesis is that if forecasts from model 1 encompass the forecasts from model 2, then the covariance between e_{1t} and e_{1t}-e_{2t} will be negative or zero (e_{1t} and e_{2t} are the two sets of forecast errors from model 1 and model 2 respectively). The alternative hypothesis is that the forecasts from model 1 do not encompass those from model 2, in which case the covariance between e_{1t} and e_{1t}-e_{2t} will be positive. The results are based on a one-step-ahead forecast covering a 100-day out-of-sample period. Forecast quality comparisons are made among pairs of the various volatility modelling approaches listed in Table 5.4. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

Comparison of Modelling Approaches	Metric HLN
T1 v. T2	6.6743**
T1 v. T3	3.7384**
T1 v. T4	2.4824**
T1 v. T5	4.5773**
T1 v. T6	5.6130**
T1 v. T7	6.3567**
T1 v. T8	5.5171**
T1 v. T9	6.2283**
T1 v. T10	6.3236**
T1 v. T11	5.8912**
T1 v. T12	6.5314**
T1 v. T13	4.4355**
T2 v. T3	0.2926
T2 v. T4	-1.2510
T2 v. T5	0.8277
T2 v. T6	-1.6576
T2 v. T7	-0.0604
T2 v. T8	-1.7334
T2 v. T9	-0.0243
T2 v. T10	3.1102**
T2 v. T11	1.3237
T2 v. T12	3.1148**
T2 v. T13	0.4360
T3 v. T4	2.4949**
T3 v. T5	3.2483**
T3 v. T6	2.5379**
T3 v. T7	2.9598**
T3 v. T8	2.5751**
T3 v. T9	2.8921**
T3 v. T10	5.0904**
T3 v. T11	4.4601**
T3 v. T12	4.4467**
T3 v. T13	2.6710**

Table 5.7: Comparing Forecast Quality between Volatility Modelling Approaches –Forecast Encompassing Test (continued)

This table reports the results of the asymptotic "forecast encompassing" test of Harvey et al. (1998) using a 5% significance level. The test statistic is denoted by HLN. The null hypothesis is that if forecasts from model 1 encompass the forecasts from model 2, then the covariance between e_{1t} and e_{1t}-e_{2t} will be negative or zero (e_{1t} and e_{2t} are the two sets of forecast errors from model 1 and model 2 respectively). The alternative hypothesis is that the forecasts from model 1 do not encompass those from model 2, in which case the covariance between e_{1t} and e_{1t}-e_{2t} will be positive. The results are based on a one-step-ahead forecast covering a 100-day out-of-sample period. Forecast quality comparisons are made among pairs of the various volatility modelling approaches listed in Table 5.4. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

Comparison of Modelling Approaches	Metric
	HLN
T4 v. T5	4.6698**
T4 v. T6	8.3765**
T4 v. T7	5.5554**
T4 v. T8	7.9951**
T4 v. T9	5.4518**
T4 v. T10	5.9663**
T4 v. T11	5.8639**
T4 v. T12	6.1850**
T4 v. T13	4.6473**
T5 v. T6	1.7597
T5 v. T7	2.4719**
T5 v. T8	1.7261
T5 v. T9	2.4281**
T5 v. T10	4.0332**
T5 v. T11	3.3419**
T5 v. T12	3.3922**
T5 v. T13	-2.8132**
T6 v. T7	1.7701
T6 v. T8	0.1706
T6 v. T9	1.6824
T6 v. T10	4.1308**
T6 v. T11	3.6762**
T6 v. T12	4.0909**
T6 v. T13	2.5606**

Table 5.7: Comparing Forecast Quality between Volatility Modelling Approaches –Forecast Encompassing Test (continued)

This table reports the results of the asymptotic "forecast encompassing" test of Harvey et al. (1998) using a 5% significance level. The test statistic is denoted by HLN. The null hypothesis is that if forecasts from model 1 encompass the forecasts from model 2, then the covariance between e_{1t} and e_{1t}-e_{2t} will be negative or zero (e_{1t} and e_{2t} are the two sets of forecast errors from model 1 and model 2 respectively). The alternative hypothesis is that the forecasts from model 1 do not encompass those from model 2, in which case the covariance between e_{1t} and e_{1t}-e_{2t} will be positive. The results are based on a one-step-ahead forecast covering a 100-day out-of-sample period. Forecast quality comparisons are made among pairs of the various volatility modelling approaches listed in Table 5.4. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

Comparison of Modelling Approaches	Metric
	HLN
T7 v. T8	1.4962
T7 v. T9	0.2045
T7 v. T10	4.4298**
T7 v. T11	4.3939**
T7 v. T12	4.1598**
T7 v. T13	3.3099**
T8 v. T9	2.2916**
T8 v. T10	5.3902**
T8 v. T11	4.3813**
T8 v. T12	5.6594**
T8 v. T13	2.9528**
T9 v. T10	4.3430**
T9 v. T11	4.3654**
T9 v. T12	4.0654**
T9 v. T13	3.2976**
T10 v. T11	0.3963
T10 v. T12	-1.5665
T10 v. T13	-0.2136
T11 v. T12	4.4789**
T11 v. T13	-0.1001
T12 v. T13	-9.4931**

Table 5.8: Results of test for "Correct Unconditional Coverage" – Normal Distribution

This table contains the results of the test for "correct unconditional coverage" in the failure series (99% VaR estimation) of the KLCI returns based on the out-of-sample VaR forecasts produced by the eleven volatility modelling and the *RiskMetrics* approaches listed below. The first column gives the names of the approaches, the second column reports the observed failure rates, the third column gives the likelihood ratio statistic for the unconditional coverage and the fourth column reports the probability of success with the null hypothesis, $H_0 : \hat{f} = \alpha$, equation (5.9) for 99% VaR coverage. The likelihood ratio statistic is given by: $\mathrm{LR_{uc}} = -2\ln\{L(p)/L(\hat{\pi})\}$. The variables are defined in the main text. The significance of these tests is denoted by ** (1% significance).

	Volatility Modelling Approach	Observed $\hat{f}$	Likelihood Ratio Statistic	p-value
T1	Non-periodic GARCH model	0.0122	3.2681	0.0706
T2	Jointly-estimated full dummy version of the PGARCH model	0.0172	31.5747**	0.0000
T3	Two-step full dummy version of the PGARCH model	0.0341	265.9433**	0.0000
T4	Jointly-estimated partial dummy version of the PGARCH model	0.0255	125.2580**	0.0000
T5	Two-step partial dummy version of the PGARCH model	0.0228	90.3686**	0.0000
T6	Jointly estimated FFF version of the PGARCH model	0.0161	23.3407**	0.0001
T7	Two-step FFF version of the PGARCH model	0.0173	32.6781**	0.0000
T8	Jointly estimated Augmented FFF version of the PGARCH model	0.0166	27.3256**	0.0000
T9	Two-step Augmented FFF version of the PGARCH model	0.0164	25.2986**	0.0000
T10	Jointly estimated Spline version of the PGARCH model	0.0162	24.3117**	0.0000
T11	Two-step Spline version of the PGARCH model	0.0165	26.3044**	0.0000
T12	Jointly estimated Augmented Spline version of the PGARCH model	0.0162	24.3117**	0.0000
T13	Two-step Augmented Spline version of the PGARCH model	0.0209	68.0997**	0.0000
	RiskMetrics	0.0296	181.8108**	0.0000

Table 5.9: Results of test for "Correct Unconditional Coverage" – Normal Distribution

This table contains the results of the test for "correct unconditional coverage" in the failure series (95% VaR estimation) of the KLCI returns based on the out-of-sample VaR forecasts produced by the eleven volatility modelling and the *RiskMetrics* approaches listed below. The first column gives the names of the approaches, the second column reports the observed failure rates, the third column gives the likelihood ratio statistic for the unconditional coverage and the fourth column reports the probability of success with the null hypothesis, $H_0 : \hat{f} = \alpha$, equation (5.9) for 95% VaR coverage. The likelihood ratio statistic is given by: $LR_{uc} = -2\ln\{L(p)/L(\hat{\pi})\}$. The variables are defined in the main text. The significance of these tests is denoted by * (5% significance).

	Volatility Modelling Approach	Observed $\hat{f}$	Likelihood Ratio Statistic	*p*-value
T1	Non-periodic GARCH model	0.0314	62.1116*	0.0000
T2	Jointly-estimated full dummy version of the PGARCH model	0.0447	4.4792*	0.0343
T3	Two-step full dummy version of the PGARCH model	0.0759	91.1227*	0.0000
T4	Jointly-estimated partial dummy version of the PGARCH model	0.0543	2.8369	0.0921
T5	Two-step partial dummy version of the PGARCH model	0.0564	6.0470*	0.0139
T6	Jointly estimated FFF version of the PGARCH model	0.0405	14.8621*	0.0001
T7	Two-step FFF version of the PGARCH model	0.0431	7.7461*	0.0062
T8	Jointly estimated Augmented FFF version of the PGARCH model	0.0495	0.0454	0.8312
T9	Two-step Augmented FFF version of the PGARCH model	0.0430	8.0603*	0.0045
T10	Jointly estimated Spline version of the PGARCH model	0.0443	5.2093*	0.0224
T11	Two-step Spline version of the PGARCH model	0.0477	0.8344	0.3610
T12	Jointly estimated Augmented Spline version of the PGARCH model	0.0457	2.9663	0.0850
T13	Two-step Augmented Spline version of the PGARCH model	0.0519	0.5511	0.4578
	RiskMetrics	0.0668	38.3912*	0.0000

Table 5.10: Results of test for "Independence" – Normal Distribution

This table contains the results of the test for "independence" in the failure series (99% VaR) estimation of the KLCI returns based on the out-of-sample VaR forecasts produced by the eleven volatility modelling and the *RiskMetrics* approaches listed below. For each approach, an OLS regression as given in equation (5.18) is performed. The first column gives the details of the approaches, the second column gives the estimated F-statistics of the hypothesis specified in (5.19) and the final column reports the corresponding *p*-values. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

	Volatility Modelling Approach	F-statistic	*p*-value
T1	Non-periodic GARCH model	9.0517**	0.0026
T2	Jointly-estimated full dummy version of the PGARCH model	10.4495**	0.0012
T3	Two-step full dummy version of the PGARCH model	13.2486**	0.0003
T4	Jointly-estimated partial dummy version of the PGARCH model	55.8022**	0.0000
T5	Two-step partial dummy version of the PGARCH model	18.3282**	0.0000
T6	Jointly estimated FFF version of the PGARCH model	18.1697**	0.0000
T7	Two-step FFF version of the PGARCH model	12.2878**	0.0006
T8	Jointly estimated Augmented FFF version of the PGARCH model	7.1667**	0.0074
T9	Two-step Augmented FFF version of the PGARCH model	6.4315**	0.0085
T10	Jointly estimated Spline version of the PGARCH model	17.5973**	0.0000
T11	Two-step Spline version of the PGARCH model	6.9068**	0.0051
T12	Jointly estimated Augmented Spline version of the PGARCH model	7.5449**	0.0060
T13	Two-step Augmented Spline version of the PGARCH model	13.0958**	0.0003
	RiskMetrics	18.7099**	0.0000

Table 5.11: Results of test for "Independence" – Normal Distribution

This table contains the results of the test for "independence" in the failure series (95% VaR) estimation of the KLCI returns based on the out-of-sample VaR forecasts produced by the eleven volatility modelling and the *RiskMetrics* approaches listed below. For each approach, an OLS regression as given in equation (5.18) is performed. The first column gives the details of the approaches, the second column gives the estimated F-statistics of the hypothesis specified in (5.19) and the final column reports the corresponding *p*-values. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

	Volatility Modelling Approach	F-statistic	*p*-value
T1	Non-periodic GARCH model	0.0288	0.8651
T2	Jointly-estimated full dummy version of the PGARCH model	2.9972	0.0835
T3	Two-step full dummy version of the PGARCH model	29.3056**	0.0000
T4	Jointly-estimated partial dummy version of the PGARCH model	49.3478**	0.0000
T5	Two-step partial dummy version of the PGARCH model	16.5607**	0.0000
T6	Jointly estimated FFF version of the PGARCH model	5.5777*	0.0182
T7	Two-step FFF version of the PGARCH model	12.6545**	0.0000
T8	Jointly estimated Augmented FFF version of the PGARCH model	9.4828**	0.0021
T9	Two-step Augmented FFF version of the PGARCH model	11.4365**	0.0000
T10	Jointly estimated Spline version of the PGARCH model	1.5230	0.2172
T11	Two-step Spline version of the PGARCH model	9.4898**	0.0014
T12	Jointly estimated Augmented Spline version of the PGARCH model	4.5239*	0.0335
T13	Two-step Augmented Spline version of the PGARCH model	20.7999**	0.0000
	RiskMetrics	28.0943**	0.0000

Table 5.12: Results of test for "Correct Conditional Coverage" – Normal Distribution

This table contains the results of the test for "correct conditional coverage" in the failure series (99% VaR) estimation of the KLCI returns based on the out-of-sample VaR forecasts produced by the eleven volatility modelling and the *RiskMetrics* approaches listed below. For each approach, an OLS regression as given in equation (5.18) is performed. The first column gives the names of the approaches, the second column gives the estimated F-statistics of the hypothesis specified in (5.17) and the final column reports the corresponding *p*-values. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

	Volatility Modelling Approach	**F-statistic**	***p*-value**
T1	Non-periodic GARCH model	5.4739**	0.0042
T2	Jointly-estimated full dummy version of the PGARCH model	14.7314**	0.0000
T3	Two-step full dummy version of the PGARCH model	66.3112**	0.0000
T4	Jointly-estimated partial dummy version of the PGARCH model	55.2764**	0.0000
T5	Two-step partial dummy version of the PGARCH model	32.4812**	0.0000
T6	Jointly estimated FFF version of the PGARCH model	15.7608**	0.0000
T7	Two-step FFF version of the PGARCH model	11.6015**	0.0000
T8	Jointly estimated Augmented FFF version of the PGARCH model	12.1847**	0.0000
T9	Two-step Augmented FFF version of the PGARCH model	9.2817**	0.0001
T10	Jointly estimated Spline version of the PGARCH model	15.7704**	0.0000
T11	Two-step Spline version of the PGARCH model	10.6209**	0.0000
T12	Jointly estimated Augmented Spline version of the PGARCH model	11.4459**	0.0000
T13	Two-step Augmented Spline version of the PGARCH model	25.2697**	0.0000
	RiskMetrics	53.0367**	0.0000

Table 5.13: Results of test for "Correct Conditional Coverage" – Normal Distribution

This table contains the results of the test for "correct conditional coverage" in the failure series (95% VaR) estimation of the KLCI returns based on the out-of-sample VaR forecasts produced by the eleven volatility modelling and the *RiskMetrics* approaches listed below. For each approach, an OLS regression as given in equation (5.18) is performed. The first column gives the names of the approaches, the second column gives the estimated F-statistics of the hypothesis specified in (5.17) and the final column reports the corresponding *p*-values. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

	Volatility Modelling Approach	F-statistic	*p*-value
T1	Non-periodic GARCH model	44.0134**	0.0000
T2	Jointly-estimated full dummy version of the PGARCH model	4.9190**	0.0000
T3	Two-step full dummy version of the PGARCH model	40.7226**	0.0000
T4	Jointly-estimated partial dummy version of the PGARCH model	24.9239**	0.0000
T5	Two-step partial dummy version of the PGARCH model	9.3301**	0.0001
T6	Jointly estimated FFF version of the PGARCH model	13.8152**	0.0000
T7	Two-step FFF version of the PGARCH model	4.7998**	0.0083
T8	Jointly estimated Augmented FFF version of the PGARCH model	5.1163**	0.0057
T9	Two-step Augmented FFF version of the PGARCH model	4.5386*	0.0107
T10	Jointly estimated Spline version of the PGARCH model	4.3788*	0.0126
T11	Two-step Spline version of the PGARCH model	10.6209**	0.0000
T12	Jointly estimated Augmented Spline version of the PGARCH model	4.8751**	0.0077
T13	Two-step Augmented Spline version of the PGARCH model	10.4479**	0.0000
	RiskMetrics	53.0367**	0.0000

Table 5.14: Results of test for "Correct Unconditional Coverage" – Student t Distribution (4 Degrees of Freedom)

This table contains the results of the test for "correct unconditional coverage" in the failure series (99% VaR estimation) of the KLCI returns based on the out-of-sample VaR forecasts produced by the eleven volatility modelling and the *RiskMetrics* approaches listed below. The first column gives the names of the approaches, the second column reports the observed failure rates, the third column gives the likelihood ratio statistic for the unconditional coverage and the fourth column reports the probability of success with the null hypothesis, $H_0 : \hat{f} = \alpha$, equation (5.9) for 99% VaR coverage. The likelihood ratio statistic is given by: $LR_{uc} = -2\ln\{L(p)/L(\hat{\pi})\}$. The variables are defined in the main text. The significance of these tests is denoted by ** (1% significance).

	Volatility Modelling Approach	Observed $\hat{f}$	Likelihood Ratio Statistic	p-value
T1	Non-periodic GARCH model	0.0028	53.4779**	0.0000
T2	Jointly-estimated full dummy version of the PGARCH model	0.0015	84.6036**	0.0000
T3	Two-step full dummy version of the PGARCH model	0.0034	44.0665**	0.0000
T4	Jointly-estimated partial dummy version of the PGARCH model	0.0054	18.9427**	0.0000
T5	Two-step partial dummy version of the PGARCH model	0.0018	77.2817**	0.0000
T6	Jointly estimated FFF version of the PGARCH model	0.0023	64.4334**	0.0000
T7	Two-step FFF version of the PGARCH model	0.0020	70.5907**	0.0000
T8	Jointly estimated Augmented FFF version of the PGARCH model	0.0026	58.7431**	0.0000
T9	Two-step Augmented FFF version of the PGARCH model	0.0023	64.4334**	0.0000
T10	Jointly estimated Spline version of the PGARCH model	0.0014	88.5152**	0.0000
T11	Two-step Spline version of the PGARCH model	0.0012	92.6503**	0.0000
T12	Jointly estimated Augmented Spline version of the PGARCH model	0.0014	88.5152**	0.0000
T13	Two-step Augmented Spline version of the PGARCH model	0.0018	77.2817**	0.0000
	RiskMetrics	0.0047	24.3177**	0.0000

Table 5.15: Results of test for "Correct Unconditional Coverage" – Student t Distribution (4 Degrees of Freedom)

This table contains the results of the test for "correct unconditional coverage" in the failure series (95% VaR estimation) of the KLCI returns based on the out-of-sample VaR forecasts produced by the eleven volatility modelling and the *RiskMetrics* approaches listed below. The first column gives the names of the approaches, the second column reports the observed failure rates, the third column gives the likelihood ratio statistic for the unconditional coverage and the fourth column reports the probability of success with the null hypothesis, $H_0 : \hat{f} = \alpha$, equation (5.9) for 95% VaR coverage. The likelihood ratio statistic is given by: $LR_{uc} = -2\ln\{L(p)/L(\hat{\pi})\}$. The variables are defined in the main text. The significance of these tests is denoted by * (5% significance).

	Volatility Modelling Approach	Observed $\hat{f}$	Likelihood Ratio Statistic	*p*-value
T1	Non-periodic GARCH model	0.0091	389.8936*	0.0000
T2	Jointly-estimated full dummy version of the PGARCH model	0.0118	325.3299*	0.0000
T3	Two-step full dummy version of the PGARCH model	0.0268	100.5804*	0.0000
T4	Jointly-estimated partial dummy version of the PGARCH model	0.0209	166.7845*	0.0000
T5	Two-step partial dummy version of the PGARCH model	0.0169	227.1436*	0.0000
T6	Jointly estimated FFF version of the PGARCH model	0.0111	340.5279*	0.0000
T7	Two-step FFF version of the PGARCH model	0.0122	316.5413*	0.0000
T8	Jointly estimated Augmented FFF version of the PGARCH model	0.0127	305.0976*	0.0000
T9	Two-step Augmented FFF version of the PGARCH model	0.0122	316.5413*	0.0000
T10	Jointly estimated Spline version of the PGARCH model	0.0127	305.0976*	0.0000
T11	Two-step Spline version of the PGARCH model	0.0130	299.5051*	0.0000
T12	Jointly estimated Augmented Spline version of the PGARCH model	0.0130	299.5051*	0.0000
T13	Two-step Augmented Spline version of the PGARCH model	0.0155	250.3677*	0.0000
	RiskMetrics	0.0222	144.6096*	0.0000

Table 5.16: Results of test for "Independence" – Student t Distribution (4 Degrees of Freedom)

This table contains the results of the test for "independence" in the failure series (99% VaR) estimation of the KLCI returns based on the out-of-sample VaR forecasts produced by the eleven volatility modelling and the *RiskMetrics* approaches listed below. For each approach, an OLS regression as given in equation (5.18) is performed. The first column gives the details of the approaches, the second column gives the estimated F-statistics of the hypothesis specified in (5.19) and the final column reports the corresponding *p*-values. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

	Volatility Modelling Approach	F-statistic	*p*-value
T1	Non-periodic GARCH model	0.7418	0.3891
T2	Jointly-estimated full dummy version of the PGARCH model	0.7408	0.3894
T3	Two-step full dummy version of the PGARCH model	0.7422	0.3890
T4	Jointly-estimated partial dummy version of the PGARCH model	32.6408**	0.0000
T5	Two-step partial dummy version of the PGARCH model	0.7410	0.3894
T6	Jointly estimated FFF version of the PGARCH model	17.7300**	0.0000
T7	Two-step FFF version of the PGARCH model	0.7412	0.3893
T8	Jointly estimated Augmented FFF version of the PGARCH model	0.7416	0.3892
T9	Two-step Augmented FFF version of the PGARCH model	0.7414	0.3891
T10	Jointly estimated Spline version of the PGARCH model	0.7407	0.3895
T11	Two-step Spline version of the PGARCH model	0.7407	0.3895
T12	Jointly estimated Augmented Spline version of the PGARCH model	0.7407	0.3895
T13	Two-step Augmented Spline version of the PGARCH model	0.7410	0.3894
	RiskMetrics	16.5635**	0.0000

Table 5.17: Results of test for "Independence" – Student t Distribution (4 Degrees of Freedom)

This table contains the results of the test for "independence" in the failure series (95% VaR) estimation of the KLCI returns based on the out-of-sample VaR forecasts produced by the eleven volatility modelling and the *RiskMetrics* approaches listed below. For each approach, an OLS regression as given in equation (5.18) is performed. The first column gives the details of the approaches, the second column gives the estimated F-statistics of the hypothesis specified in (5.19) and the final column reports the corresponding *p*-values. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

	Volatility Modelling Approach	F-statistic	*p*-value
T1	Non-periodic GARCH model	0.1521	0.6965
T2	Jointly-estimated full dummy version of the PGARCH model	0.1213	0.7277
T3	Two-step full dummy version of the PGARCH model	0.1571	0.6919
T4	Jointly-estimated partial dummy version of the PGARCH model	8.9282**	0.0028
T5	Two-step partial dummy version of the PGARCH model	0.2713	0.6025
T6	Jointly estimated FFF version of the PGARCH model	1.2424	0.2650
T7	Two-step FFF version of the PGARCH model	5.6119*	0.0179
T8	Jointly estimated Augmented FFF version of the PGARCH model	0.3624	0.5472
T9	Two-step Augmented FFF version of the PGARCH model	5.6119*	0.0179
T10	Jointly estimated Spline version of the PGARCH model	1.4368	0.2307
T11	Two-step Spline version of the PGARCH model	6.3757*	0.0116
T12	Jointly estimated Augmented Spline version of the PGARCH model	0.0325	0.8568
T13	Two-step Augmented Spline version of the PGARCH model	2.8849	0.0895
	RiskMetrics	0.0064	0.9364

Table 5.18: Results of test for "Correct Conditional Coverage" – Student t Distribution (4 Degrees of Freedom)

This table contains the results of the test for "correct conditional coverage" in the failure series (99% VaR) estimation of the KLCI returns based on the out-of-sample VaR forecasts produced by the eleven volatility modelling and the *RiskMetrics* approaches listed below. For each approach, an OLS regression as given in equation (5.18) is performed. The first column gives the names of the approaches, the second column gives the estimated F-statistics of the hypothesis specified in (5.17) and the final column reports the corresponding p-values. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

	Volatility Modelling Approach	**F-statistic**	***p*-value**
T1	Non-periodic GARCH model	67.0584**	0.0000
T2	Jointly-estimated full dummy version of the PGARCH model	180.5801**	0.0000
T3	Two-step full dummy version of the PGARCH model	48.1932**	0.0000
T4	Jointly-estimated partial dummy version of the PGARCH model	33.3263**	0.0000
T5	Two-step partial dummy version of the PGARCH model	143.2934**	0.0000
T6	Jointly estimated FFF version of the PGARCH model	107.9377**	0.0000
T7	Two-step FFF version of the PGARCH model	116.2140**	0.0000
T8	Jointly estimated Augmented FFF version of the PGARCH model	79.7874**	0.0000
T9	Two-step Augmented FFF version of the PGARCH model	95.7395**	0.0000
T10	Jointly estimated Spline version of the PGARCH model	204.9468**	0.0000
T11	Two-step Spline version of the PGARCH model	234.8781**	0.0000
T12	Jointly estimated Augmented Spline version of the PGARCH model	204.9648**	0.0000
T13	Two-step Augmented Spline version of the PGARCH model	143.2934**	0.0000
	RiskMetrics	32.1608**	0.0000

Table 5.19: Results of test for "Correct Conditional Coverage" – Student t Distribution (4 Degrees of Freedom)

This table contains the results of the test for "correct conditional coverage" in the failure series (95% VaR) estimation of the KLCI returns based on the out-of-sample VaR forecasts produced by the eleven volatility modelling and the *RiskMetrics* approaches listed below. For each approach, an OLS regression as given in equation (5.18) is performed. The first column gives the names of the approaches, the second column gives the estimated F-statistics of the hypothesis specified in (5.17) and the final column reports the corresponding p-values. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

	Volatility Modelling Approach	F-statistic	p-value
T1	Non-periodic GARCH model	696.8286**	0.0000
T2	Jointly-estimated full dummy version of the PGARCH model	470.4661**	0.0000
T3	Two-step full dummy version of the PGARCH model	78.0893**	0.0000
T4	Jointly-estimated partial dummy version of the PGARCH model	168.3339**	0.0000
T5	Two-step partial dummy version of the PGARCH model	250.3084**	0.0000
T6	Jointly estimated FFF version of the PGARCH model	513.9741**	0.0000
T7	Two-step FFF version of the PGARCH model	441.0879**	0.0000
T8	Jointly estimated Augmented FFF version of the PGARCH model	413.9792**	0.0000
T9	Two-step Augmented FFF version of the PGARCH model	441.0879**	0.0000
T10	Jointly estimated Spline version of the PGARCH model	421.1928**	0.0000
T11	Two-step Spline version of the PGARCH model	396.1689**	0.0000
T12	Jointly estimated Augmented Spline version of the PGARCH model	402.3475**	0.0000
T13	Two-step Augmented Spline version of the PGARCH model	299.0286**	0.0000
	RiskMetrics	135.0970**	0.0000

Table 5.20: Results of test for "Correct Unconditional Coverage" – Student t Distribution (24 Degrees of Freedom)

This table contains the results of the test for "correct unconditional coverage" in the failure series (99% VaR estimation) of the KLCI returns based on the out-of-sample VaR forecasts produced by the eleven volatility modelling and the *RiskMetrics* approaches listed below. The first column gives the names of the approaches, the second column reports the observed failure rates, the third column gives the likelihood ratio statistic for the unconditional coverage and the fourth column reports the probability of success with the null hypothesis, $H_0 : \hat{f} = \alpha$, equation (5.9) for 99% VaR coverage. The likelihood ratio statistic is given by: $\mathrm{LR}_{uc} = -2\ln\{L(p)/L(\hat{\pi})\}$. The variables are defined in the main text. The significance of these tests is denoted by ** (1% significance).

	Volatility Modelling Approach	Observed $\hat{f}$	Likelihood Ratio Statistic	p-value
T1	Non-periodic GARCH model	0.0089	0.9063	0.3410
T2	Jointly-estimated full dummy version of the PGARCH model	0.0118	2.1826	0.1395
T3	Two-step full dummy version of the PGARCH model	0.0257	128.1708**	0.0000
T4	Jointly-estimated partial dummy version of the PGARCH model	0.0205	63.6567**	0.0000
T5	Two-step partial dummy version of the PGARCH model	0.0166	27.3256**	0.0000
T6	Jointly estimated FFF version of the PGARCH model	0.0104	0.1212	0.7277
T7	Two-step FFF version of the PGARCH model	0.0120	2.8840	0.0894
T8	Jointly estimated Augmented FFF version of the PGARCH model	0.0126	4.5561	0.0328
T9	Two-step Augmented FFF version of the PGARCH model	0.0120	2.8840	0.0894
T10	Jointly estimated Spline version of the PGARCH model	0.0124	4.1047	0.0427
T11	Two-step Spline version of the PGARCH model	0.0130	6.0401	0.0139
T12	Jointly estimated Augmented Spline version of the PGARCH model	0.0128	5.5237	0.0187
T13	Two-step Augmented Spline version of the PGARCH model	0.0153	17.8785**	0.0000
	RiskMetrics	0.0219	88.3806**	0.0000

Table 5.21: Results of test for "Correct Unconditional Coverage" – Student t Distribution (24 Degrees of Freedom)

This table contains the results of the test for "correct unconditional coverage" in the failure series (95% VaR estimation) of the KLCI returns based on the out-of-sample VaR forecasts produced by the eleven volatility modelling and the *RiskMetrics* approaches listed below. The first column gives the names of the approaches, the second column reports the observed failure rates, the third column gives the likelihood ratio statistic for the unconditional coverage and the fourth column reports the probability of success with the null hypothesis, $H_0 : \hat{f} = \alpha$, equation (5.9) for 95% VaR coverage. The likelihood ratio statistic is given by: $LR_{uc} = -2\ln\{L(p)/L(\hat{\pi})\}$. The variables are defined in the main text. The significance of these tests is denoted by * (5% significance).

	Volatility Modelling Approach	**Observed $\hat{f}$**	**Likelihood Ratio Statistic**	***p*-value**
T1	Non-periodic GARCH model	0.0262	105.8597*	0.0000
T2	Jointly-estimated dummy version of the PGARCH model	0.0381	23.9169*	0.0000
T3	Two-step dummy version of the PGARCH model	0.0659	36.1605*	0.0000
T4	Jointly-estimated dummy version of the PGARCH model	0.0485	0.3475	0.5555
T5	Two-step dummy version of the PGARCH model	0.0495	0.0454	0.8312
T6	Jointly estimated FFF version of the PGARCH model	0.0365	31.2720*	0.0000
T7	Two-step FFF version of the PGARCH model	0.0368	29.9701*	0.0000
T8	Jointly estimated Augmented FFF version of the PGARCH model	0.0426	9.0429*	0.0026
T9	Two-step Augmented FFF version of the PGARCH model	0.0369	29.3312*	0.0000
T10	Jointly estimated Spline version of the PGARCH model	0.0374	26.8497*	0.0000
T11	Two-step Spline version of the PGARCH model	0.0407	14.4265*	0.0001
T12	Jointly estimated Augmented Spline version of the PGARCH model	0.0377	25.6544*	0.0000
T13	Two-step Augmented Spline version of the PGARCH model	0.0450	4.0243*	0.0448
	RiskMetrics	0.0576	10.2550*	0.0013

Table 5.22: Results of test for "Independence" – Student t Distribution (24 Degrees of Freedom)

This table contains the results of the test for "independence" in the failure series (99% VaR) estimation of the KLCI returns based on the out-of-sample VaR forecasts produced by the eleven volatility modelling approaches and the *RiskMetrics* listed below. For each approach, an OLS regression as given in equation (5.18) is performed. The first column gives the details of the approaches, the second column gives the estimated F-statistics of the hypothesis specified in (5.19) and the final column reports the corresponding *p*-values. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

	Volatility Modelling Approach	F-statistic	*p*-value
T1	Non-periodic GARCH model	9.6212**	0.0019
T2	Jointly-estimated full dummy version of the PGARCH model	9.6998**	0.0018
T3	Two-step full dummy version of the PGARCH model	10.6067**	0.0011
T4	Jointly-estimated partial dummy version of the PGARCH model	36.3383**	0.0000
T5	Two-step partial dummy version of the PGARCH model	16.5809**	0.0000
T6	Jointly estimated FFF version of the PGARCH model	6.4145*	0.0113
T7	Two-step FFF version of the PGARCH model	21.2647**	0.0000
T8	Jointly estimated Augmented FFF version of the PGARCH model	8.3138**	0.0039
T9	Two-step Augmented FFF version of the PGARCH model	17.2993**	0.0000
T10	Jointly estimated Spline version of the PGARCH model	22.9025**	0.0000
T11	Two-step Spline version of the PGARCH model	26.8896**	0.0000
T12	Jointly estimated Augmented Spline version of the PGARCH model	13.6407**	0.0002
T13	Two-step Augmented Spline version of the PGARCH model	20.3159**	0.0000
	RiskMetrics	11.8489**	0.0006

Table 5.23: Results of test for "Independence" – Student t Distribution (24 Degrees of Freedom)

This table contains the results of the test for "independence" in the failure series (95% VaR) estimation of the KLCI returns based on the out-of-sample VaR forecasts produced by the eleven volatility modelling and the *RiskMetrics* approaches listed below. For each approach, an OLS regression as given in equation (5.18) is performed. The first column gives the details of the approaches, the second column gives the estimated F-statistics of the hypothesis specified in (5.19) and the final column reports the corresponding *p*-values. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

	Volatility Modelling Approach	F-statistic	*p*-value
T1	Non-periodic GARCH model	0.0858	0.7697
T2	Jointly-estimated full dummy version of the PGARCH model	1.4719	0.2251
T3	Two-step full dummy version of the PGARCH model	21.8269**	0.0000
T4	Jointly-estimated partial dummy version of the PGARCH model	35.3169**	0.0000
T5	Two-step partial dummy version of the PGARCH model	6.6576**	0.0099
T6	Jointly estimated FFF version of the PGARCH model	4.5419*	0.0331
T7	Two-step FFF version of the PGARCH model	3.2092	0.0733
T8	Jointly estimated Augmented FFF version of the PGARCH model	9.9944**	0.0016
T9	Two-step Augmented FFF version of the PGARCH model	14.5861**	0.0000
T10	Jointly estimated Spline version of the PGARCH model	3.7930	0.0515
T11	Two-step Spline version of the PGARCH model	11.4961**	0.0000
T12	Jointly estimated Augmented Spline version of the PGARCH model	2.5209	0.1124
T13	Two-step Augmented Spline version of the PGARCH model	4.8749*	0.0273
	RiskMetrics	20.5522**	0.0000

Table 5.24: Results of test for "Correct Conditional Coverage" – Student t Distribution (24 Degrees of Freedom)

This table contains the results of the test for "correct conditional coverage" in the failure series (99% VaR) estimation of the KLCI returns based on the out-of-sample VaR forecasts produced by the eleven volatility modelling and the *RiskMetrics* approaches listed below. For each approach, an OLS regression as given in equation (5.18) is performed. The first column gives the names of the approaches, the second column gives the estimated F-statistics of the hypothesis specified in (5.17) and the final column reports the corresponding *p*-values. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

	Volatility Modelling Approach	F-statistic	*p*-value
T1	Non-periodic GARCH model	5.6360**	0.0036
T2	Jointly-estimated full dummy version of the PGARCH model	5.4249**	0.0044
T3	Two-step full dummy version of the PGARCH model	38.1919**	0.0000
T4	Jointly-estimated partial dummy version of the PGARCH model	33.8514**	0.0000
T5	Two-step partial dummy version of the PGARCH model	16.1611**	0.0000
T6	Jointly estimated FFF version of the PGARCH model	3.2112*	0.0404
T7	Two-step FFF version of the PGARCH model	14.6798**	0.0000
T8	Jointly estimated Augmented FFF version of the PGARCH model	5.5569**	0.0039
T9	Two-step Augmented FFF version of the PGARCH model	8.7895**	0.0008
T10	Jointly estimated Spline version of the PGARCH model	12.3853**	0.0000
T11	Two-step Spline version of the PGARCH model	16.7902**	0.0000
T12	Jointly estimated Augmented Spline version of the PGARCH model	8.3760**	0.0002
T13	Two-step Augmented Spline version of the PGARCH model	15.1793**	0.0000
	RiskMetrics	27.4250**	0.0000

Table 5.25: Results of test for "Correct Conditional Coverage" – Student t Distribution (24 Degrees of Freedom)

This table contains the results of the test for "correct conditional coverage" in the failure series (95% VaR) estimation of the KLCI returns based on the out-of-sample VaR forecasts produced by the eleven volatility modelling and the *RiskMetrics* approaches listed below. For each approach, an OLS regression as given in equation (5.18) is performed. The first column gives the names of the approaches, the second column gives the estimated F-statistics of the hypothesis specified in (5.17) and the final column reports the corresponding *p*-values. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

	Volatility Modelling Approach	F-statistic	*p*-value
T1	Non-periodic GARCH model	83.5690**	0.0000
T2	Jointly-estimated full dummy version of the PGARCH model	16.9102**	0.0000
T3	Two-step full dummy version of the PGARCH model	20.9637**	0.0000
T4	Jointly-estimated partial dummy version of the PGARCH model	19.5575**	0.0000
T5	Two-step partial dummy version of the PGARCH model	3.6536*	0.0259
T6	Jointly estimated FFF version of the PGARCH model	24.9218**	0.0000
T7	Two-step FFF version of the PGARCH model	18.5138**	0.0000
T8	Jointly estimated Augmented FFF version of the PGARCH model	12.6098**	0.0000
T9	Two-step Augmented FFF version of the PGARCH model	18.3668**	0.0000
T10	Jointly estimated Spline version of the PGARCH model	21.0492**	0.0000
T11	Two-step Spline version of the PGARCH model	8.2988**	0.0000
T12	Jointly estimated Augmented Spline version of the PGARCH model	19.1093**	0.0000
T13	Two-step Augmented Spline version of the PGARCH model	5.8184**	0.0030
	RiskMetrics	11.8335**	0.0000

Figure 5.1: Forecasting Performance of Volatility Modelling Approach

The chart below compares the plots of the mean realized volatility and the forecasted return volatility (the daily average of 5-minute one-step-ahead forecasts) for a 100-day out-of-sample forecasting period using the following approach:

T1 = Non-periodic GARCH model

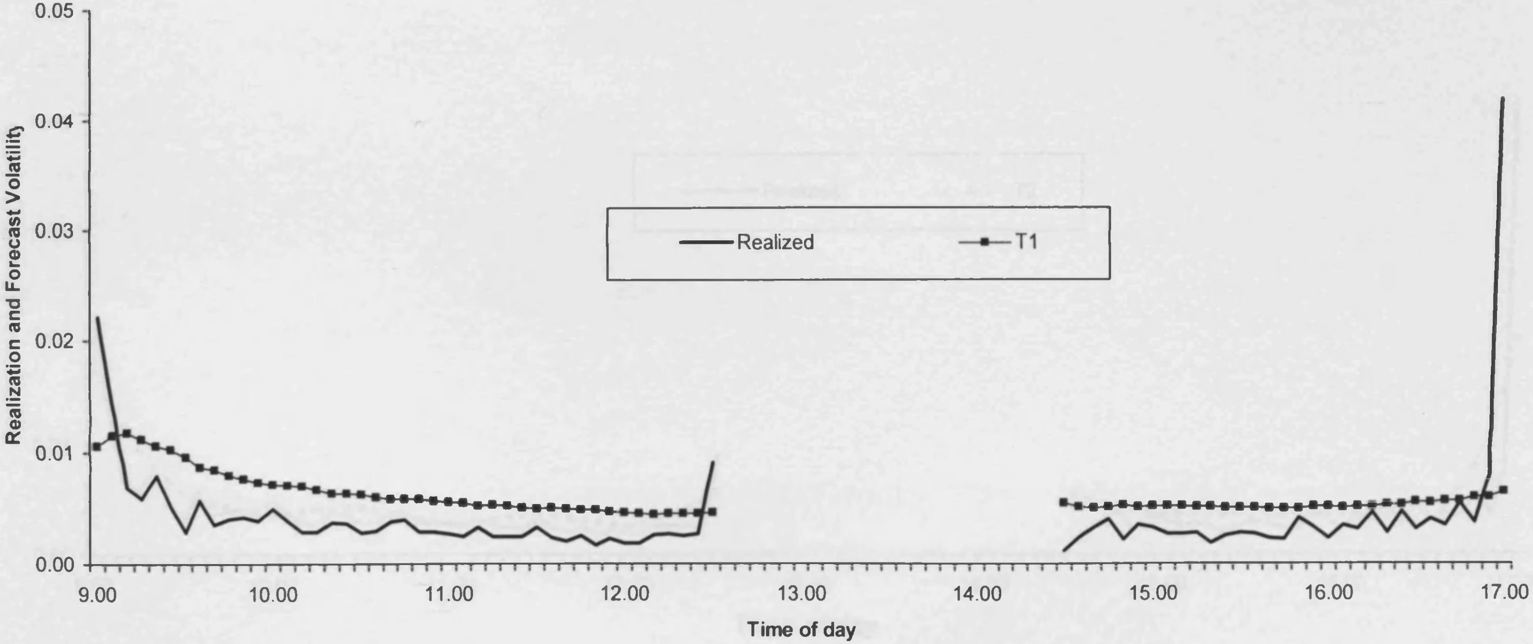


Figure 5.2: Forecasting Performance of Volatility Modelling Approach

The chart below compares the plots of the mean realized volatility and the forecasted return volatility (the daily average of 5-minute one-step-ahead forecasts) for a 100-day out-of-sample forecasting period using the following approach:

T2 = Jointly estimated full dummy version of the PGARCH model

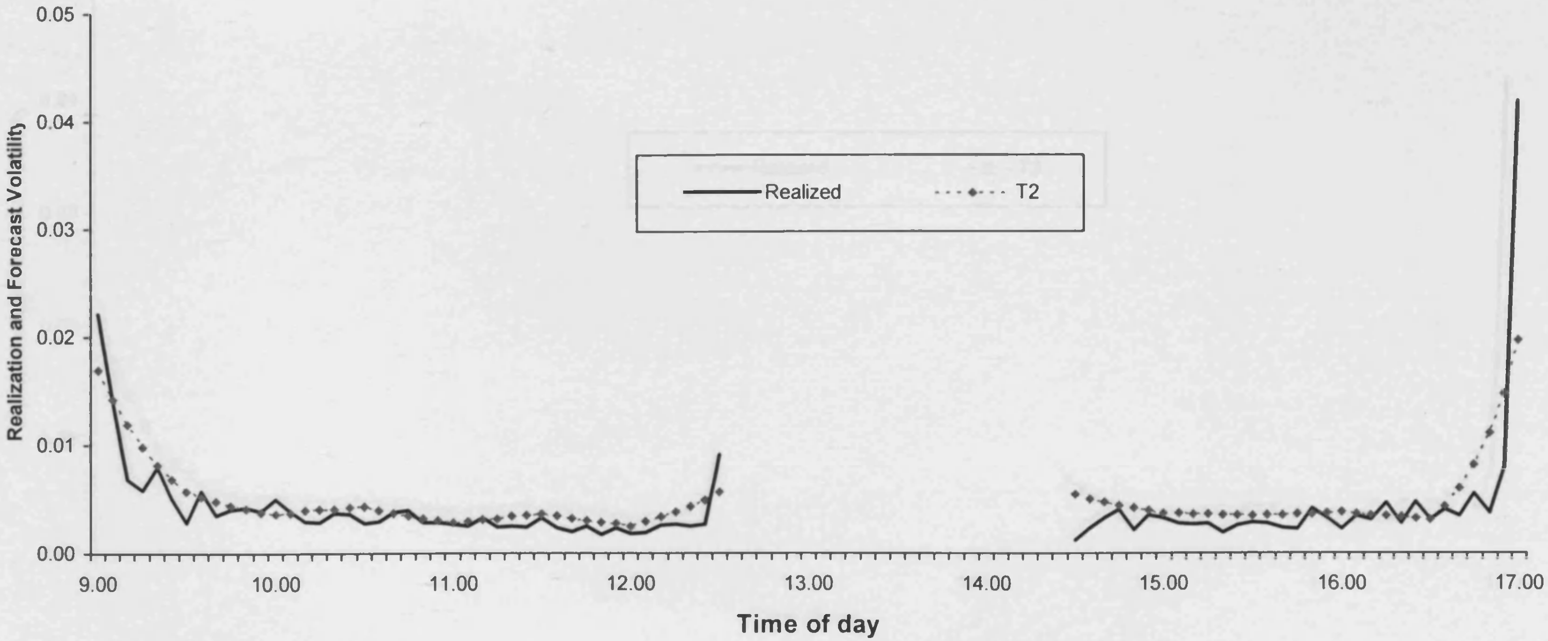


Figure 5.3: Forecasting Performance of Volatility Modelling Approach

The chart below compares the plots of the mean realized volatility and the forecasted return volatility (the daily average of 5-minute one-step-ahead forecasts) for a 100-day out-of-sample forecasting period using the following approach:

T3 = Two-step full dummy version of the PGARCH model

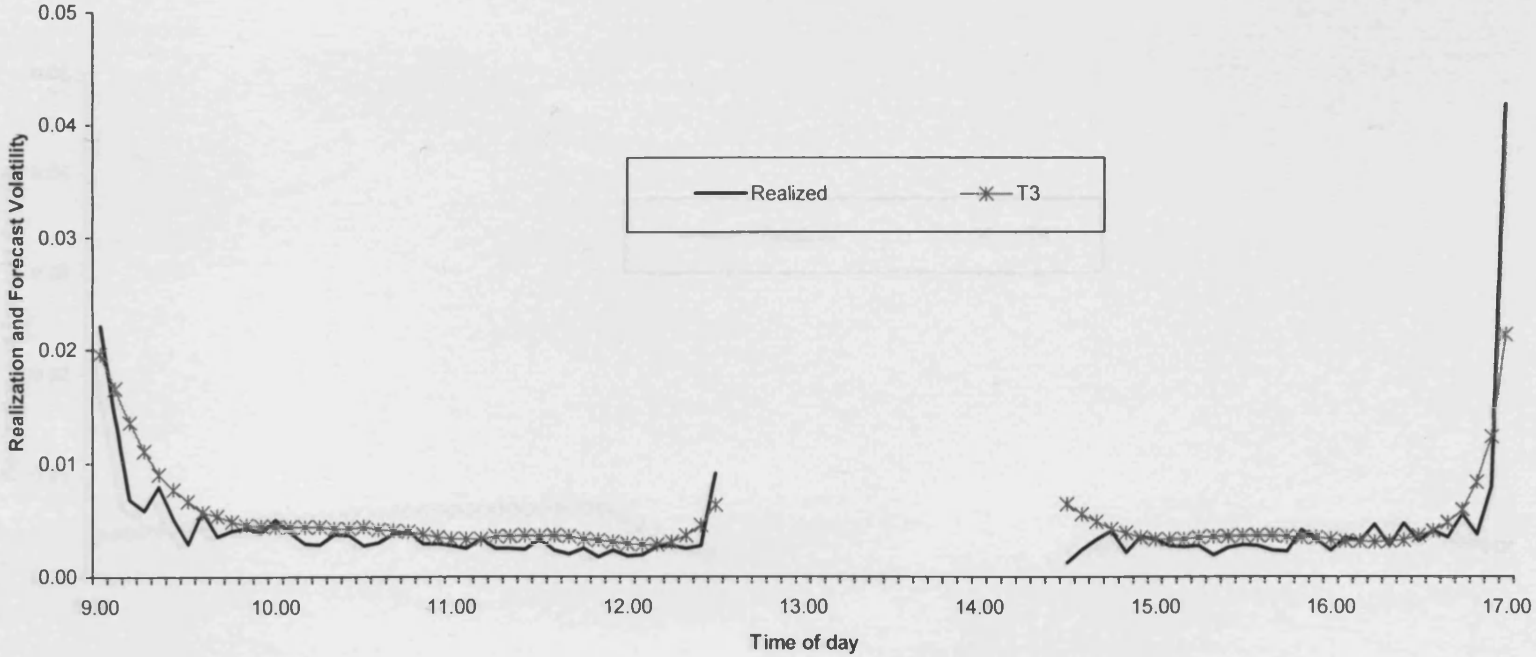


Figure 5.4: Forecasting Performance of Volatility Modelling Approach

The chart below compares the plots of the mean realized volatility and the forecasted return volatility (the daily average of 5-minute one-step-ahead forecasts) for a 100-day out-of-sample forecasting period using the following approach:

T4 = Jointly estimated partial dummy version of the PGRACH model

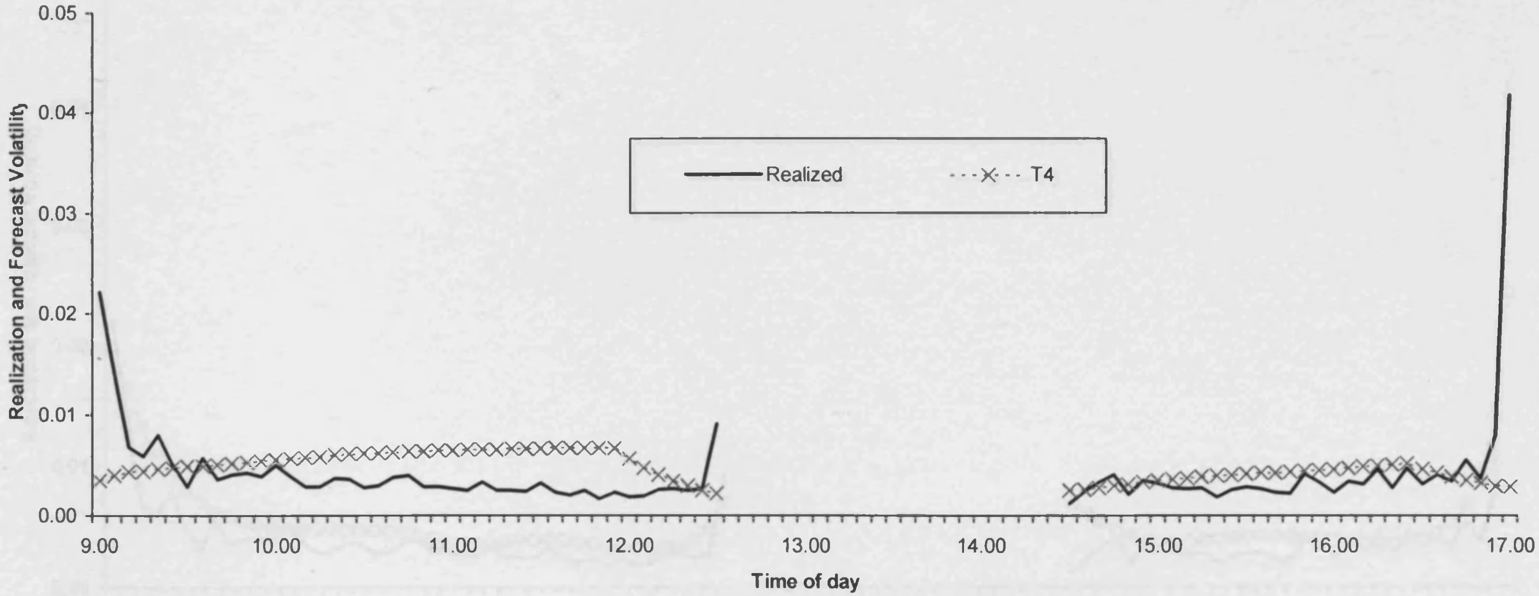


Figure 5.5: Forecasting Performance of Volatility Modelling Approach

The chart below compares the plots of the mean realized volatility and the forecasted return volatility (the daily average of 5-minute one-step-ahead forecasts) for a 100-day out-of-sample forecasting period using the following approach:

T5 = Two-step partial dummy version of the PGARCH model

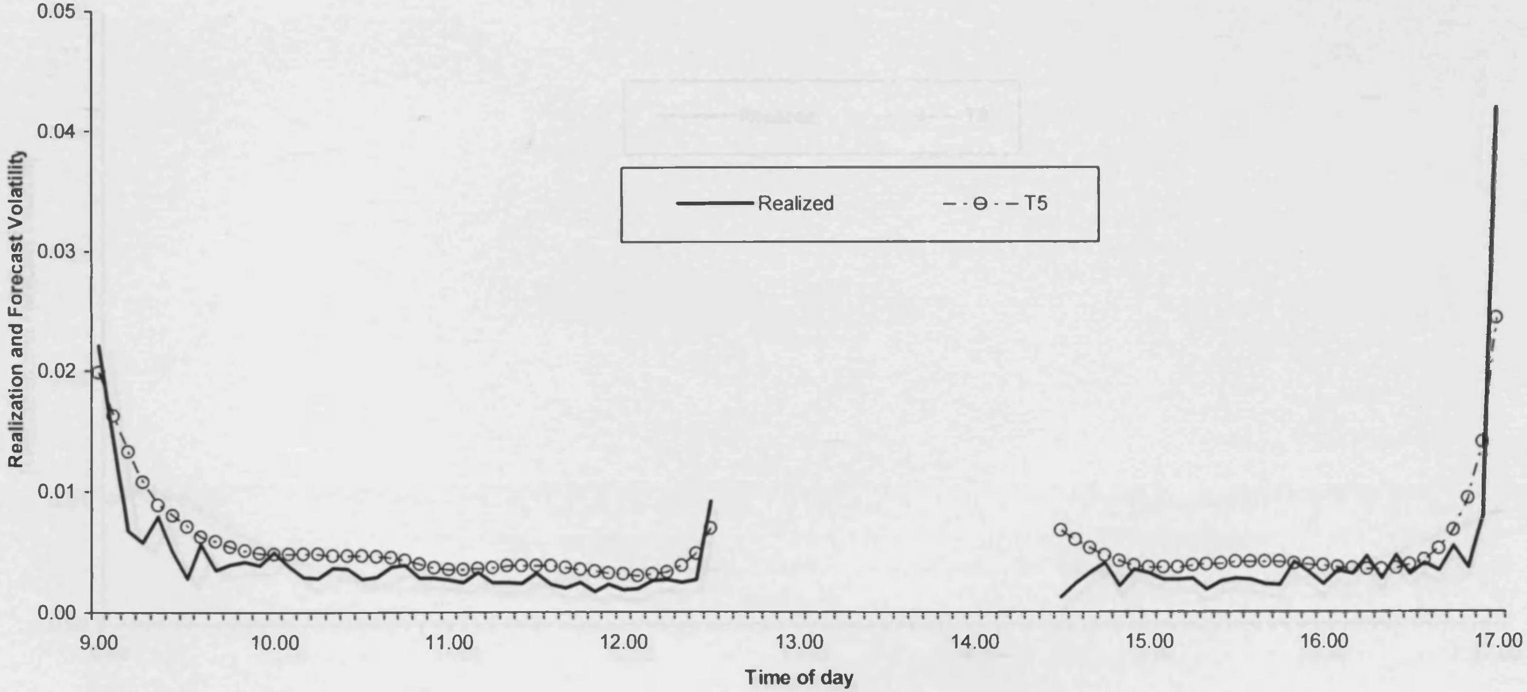


Figure 5.6: Forecasting Performance of Volatility Modelling Approach

The chart below compares the plots of the mean realized volatility and the forecasted return volatility (the daily average of 5-minute one-step-ahead forecasts) for a 100-day out-of-sample forecasting period using the following approach:

T6 = Jointly estimated FFF version of the PGARCH model

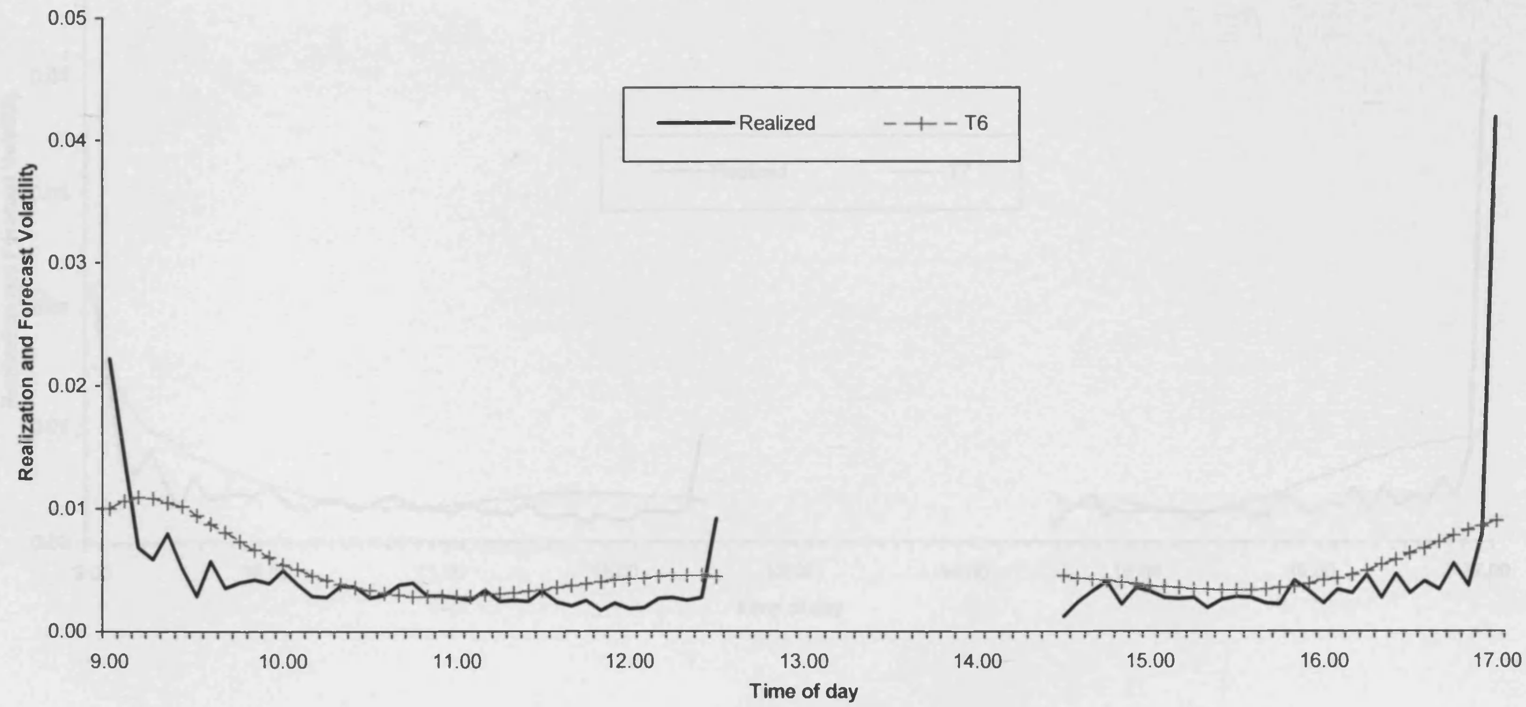


Figure 5.7: Forecasting Performance of Volatility Modelling Approach

The chart below compares the plots of the mean realized volatility and the forecasted return volatility (the daily average of 5-minute one-step-ahead forecasts) for a 100-day out-of-sample forecasting period using the following approach:

T7 = Two-step FFF version of the PGARCH model

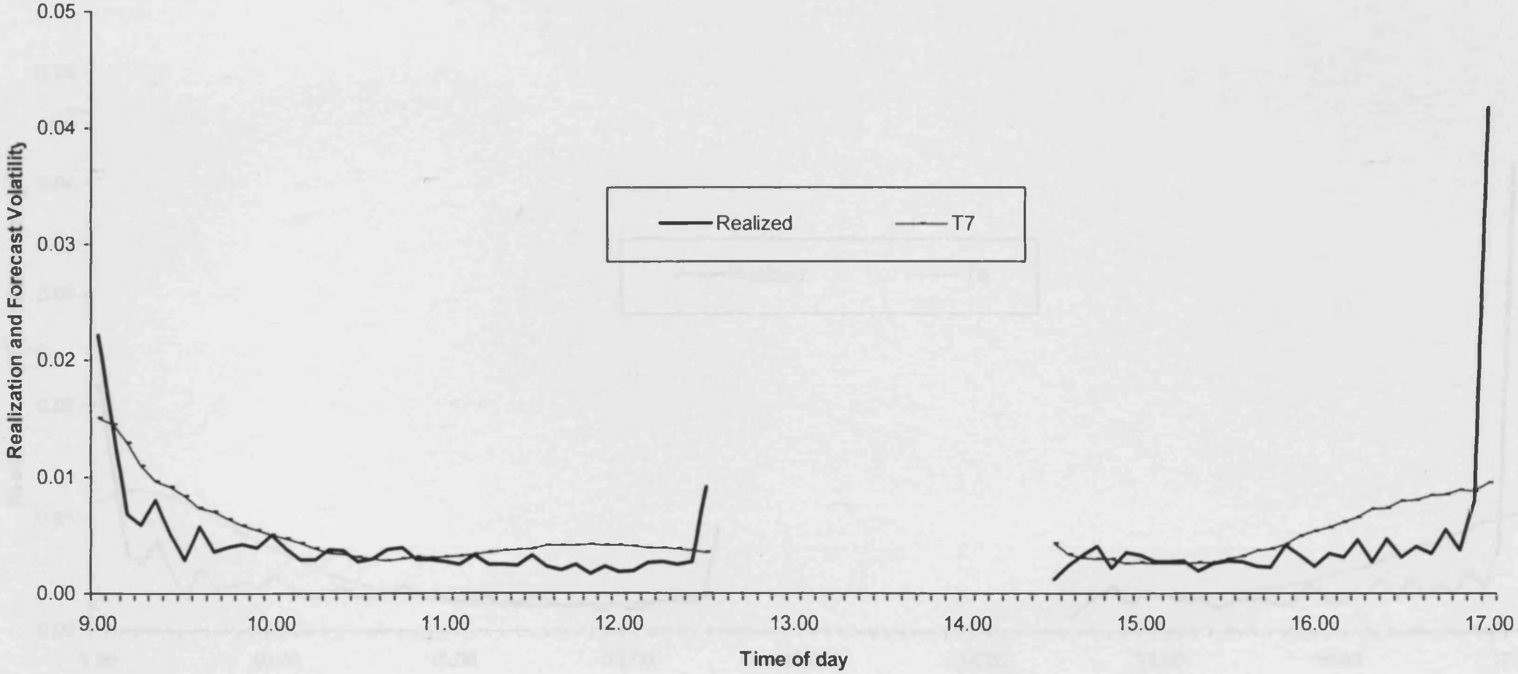


Figure 5.8: Forecasting Performance of Volatility Modelling Approach

The chart below compares the plots of the mean realized volatility and the forecasted return volatility (the daily average of 5-minute one-step-ahead forecasts) for a 100-day out-of-sample forecasting period using the following approach:

T8 = Jointly estimated Augmented FFF version of the PGARCH model

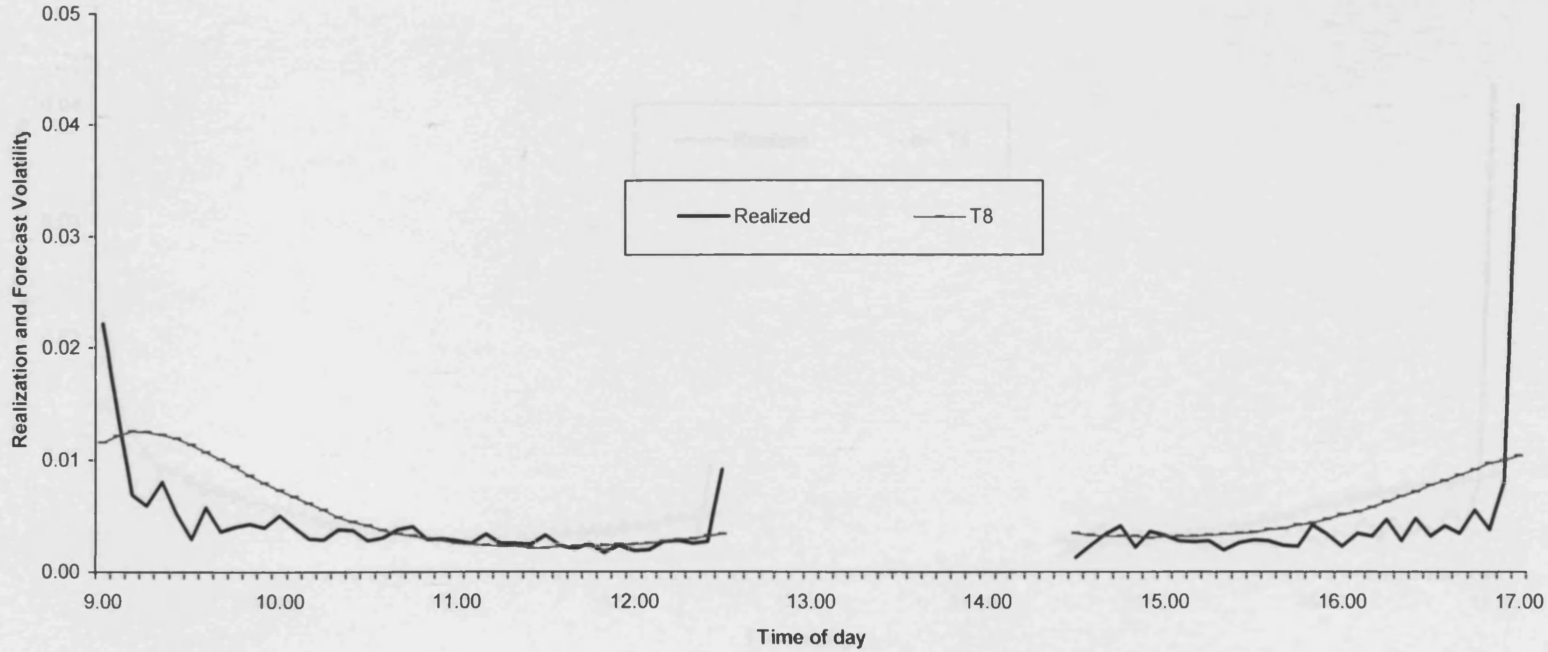


Figure 5.9: Forecasting Performance of Volatility Modelling Approach

The chart below compares the plots of the mean realized volatility and the forecasted return volatility (the daily average of 5-minute one-step-ahead forecasts) for a 100-day out-of-sample forecasting period using the following approach:

T9 = Two-step Augmented FFF version of the PGARCH model

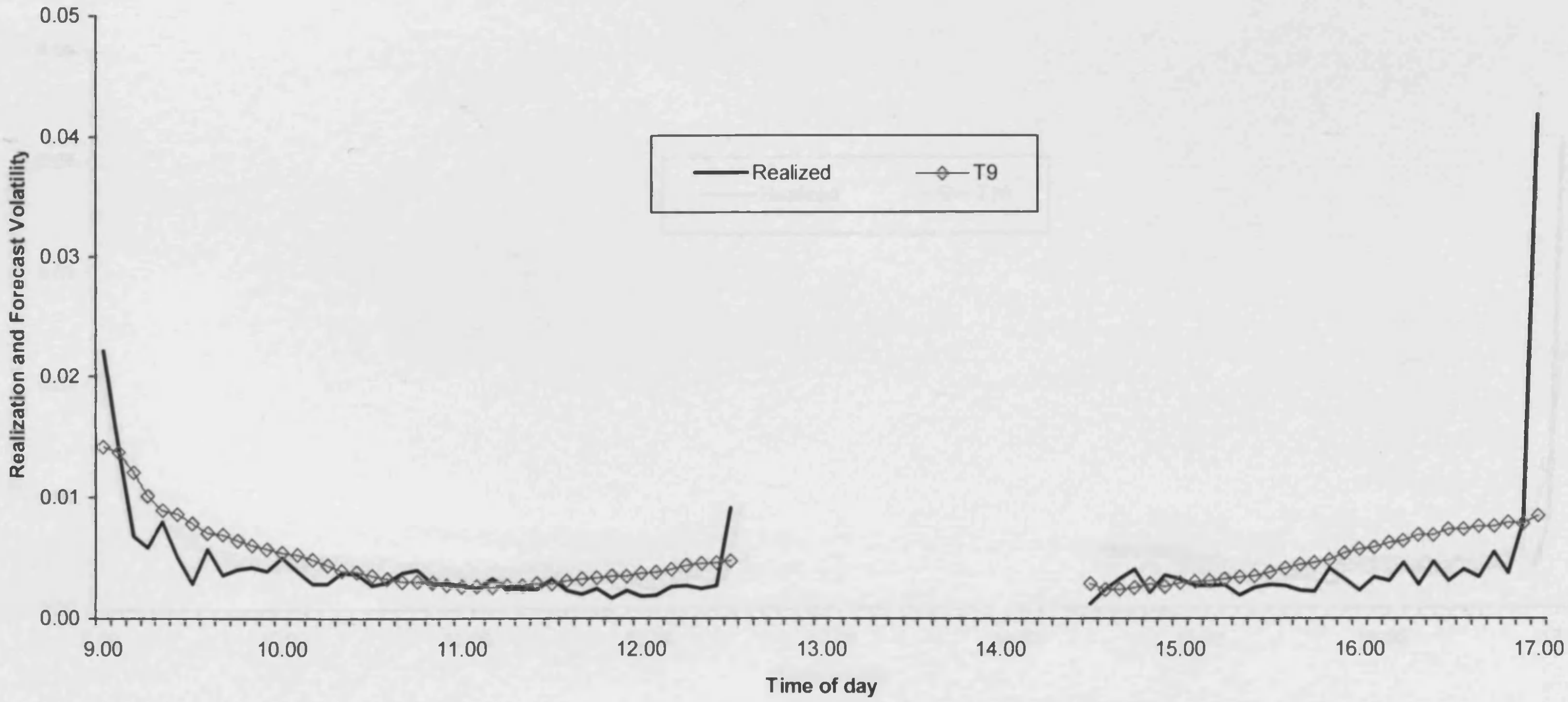


Figure 5.10: Forecasting Performance of Volatility Modelling Approach

The chart below compares the plots of the mean realized volatility and the forecasted return volatility (the daily average of 5-minute one-step-ahead forecasts) for a 100-day out-of-sample forecasting period using the following approach:

T10 = Jointly estimated Spline version of the PGARCH model

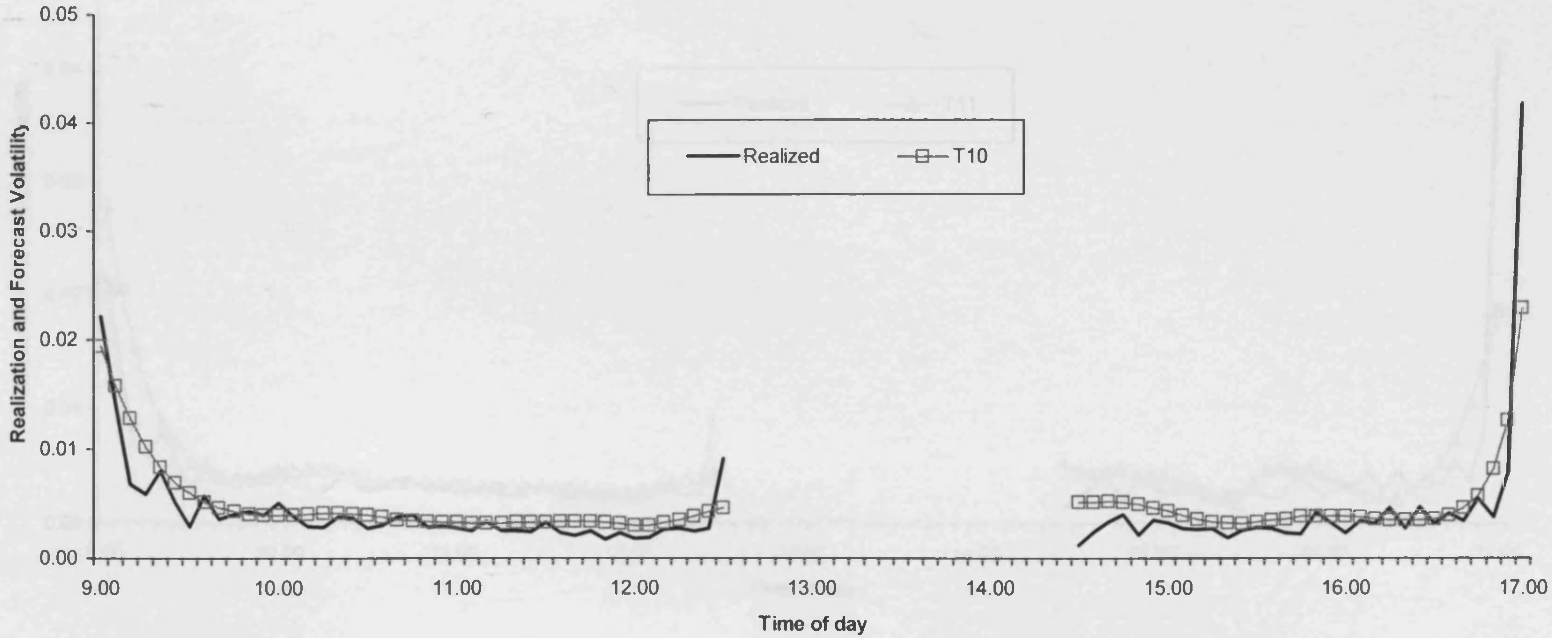


Figure 5.11: Forecasting Performance of Volatility Modelling Approach

The chart below compares the plots of the mean realized volatility and the forecasted return volatility (the daily average of 5-minute one-step-ahead forecasts) for a 100-day out-of-sample forecasting period using the following approach:

T11= Two-step Spline version of the PGARCH model

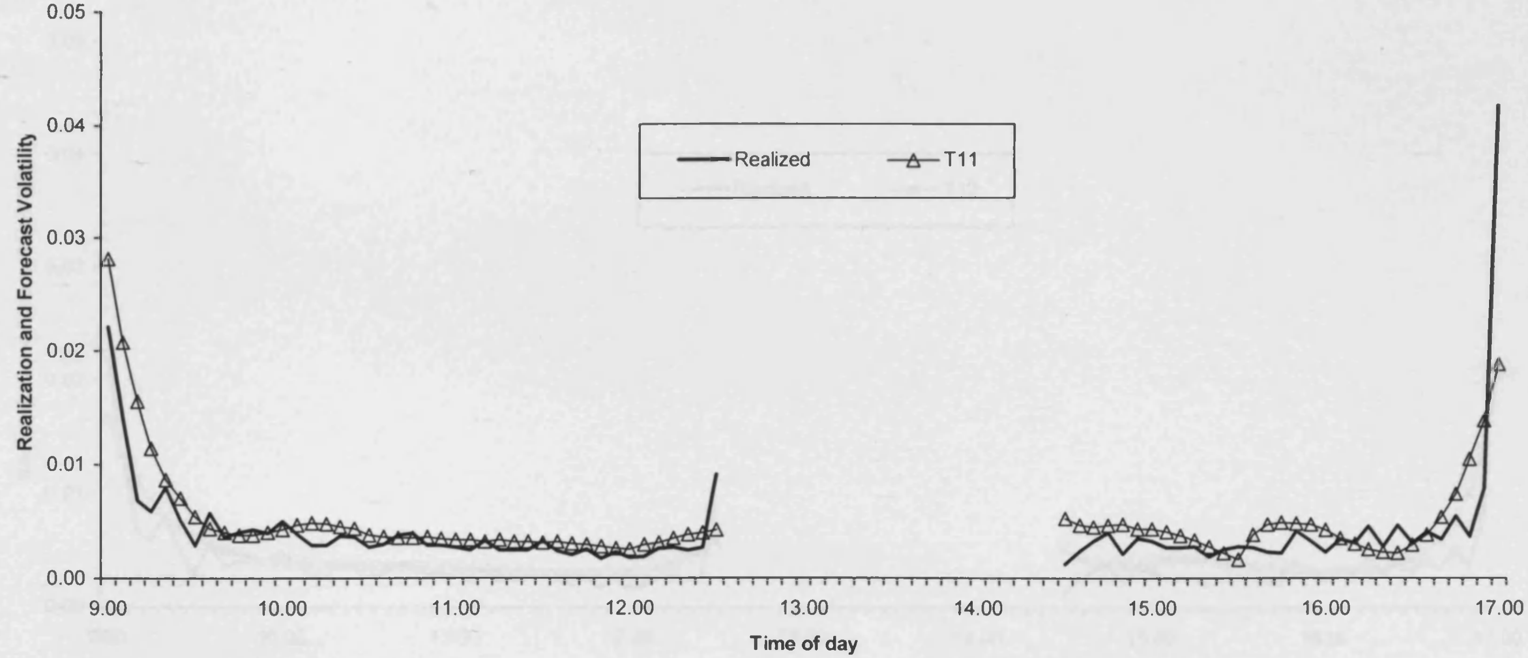


Figure 5.12: Forecasting Performance of Volatility Modelling Approach

The chart below compares the plots of the mean realized volatility and the forecasted return volatility (the daily average of 5-minute one-step-ahead forecasts) for a 100-day out-of-sample forecasting period using the following approach:

T12= Jointly estimated Augmented Spline version of the PGARCH model

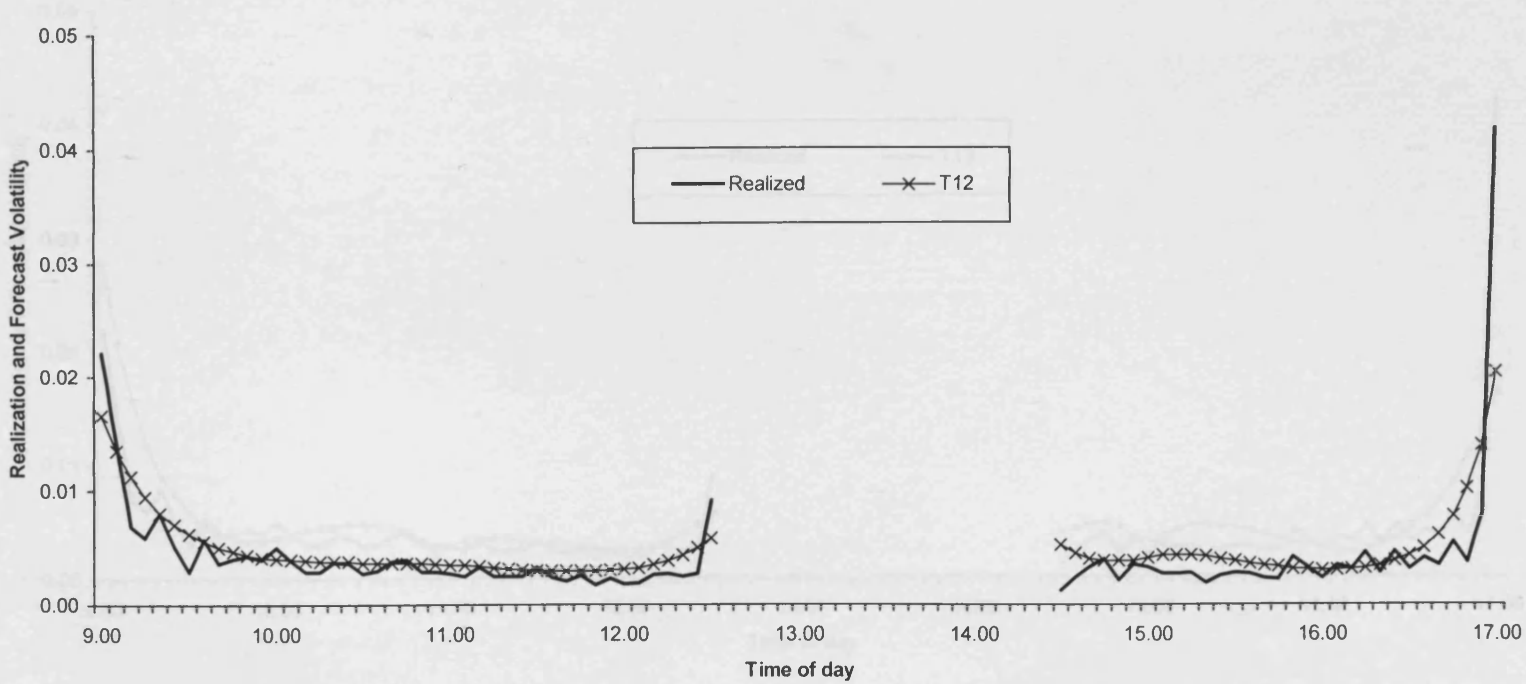


Figure 5.13: Forecasting Performance of Volatility Modelling Approach

The chart below compares the plots of the mean realized volatility and the forecasted return volatility (the daily average of 5-minute one-step-ahead forecasts) for a 100-day out-of-sample forecasting period using the following approach:

T13= Two-step Augmented Spline version of the PGARCH model

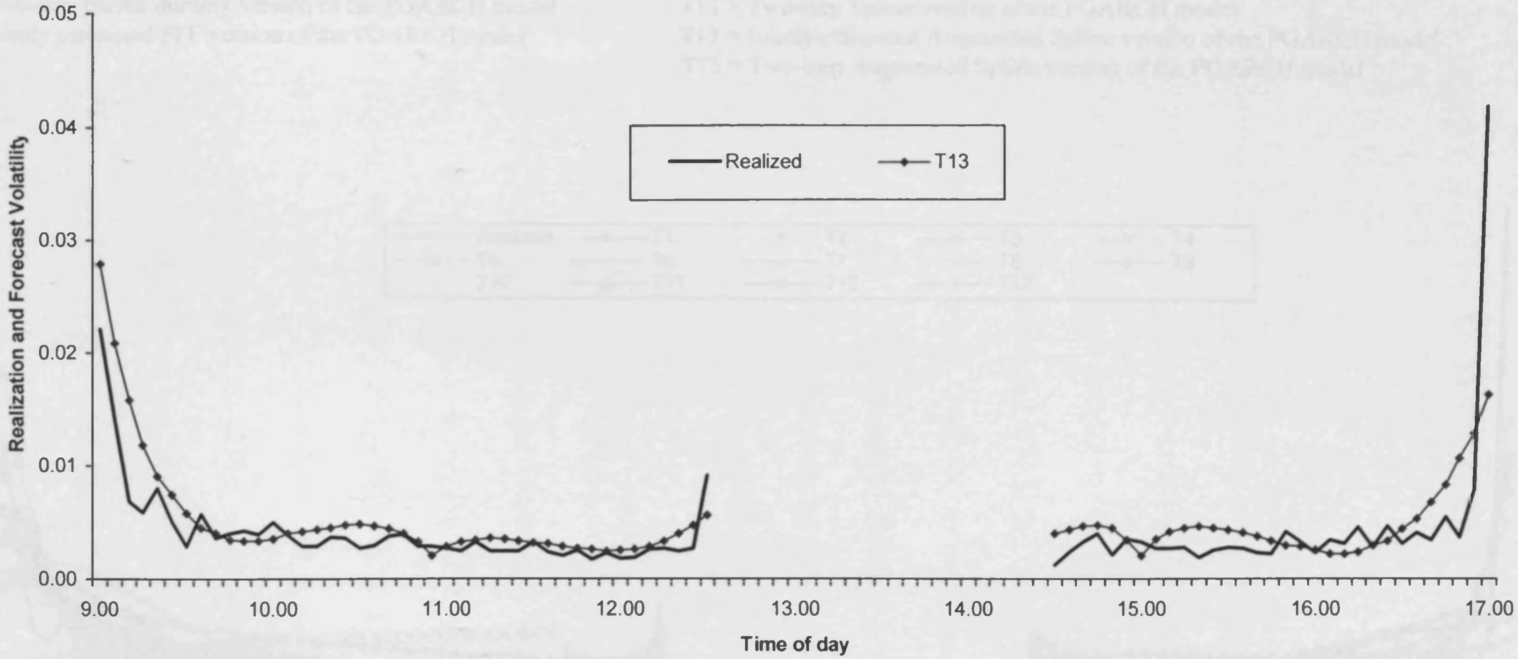
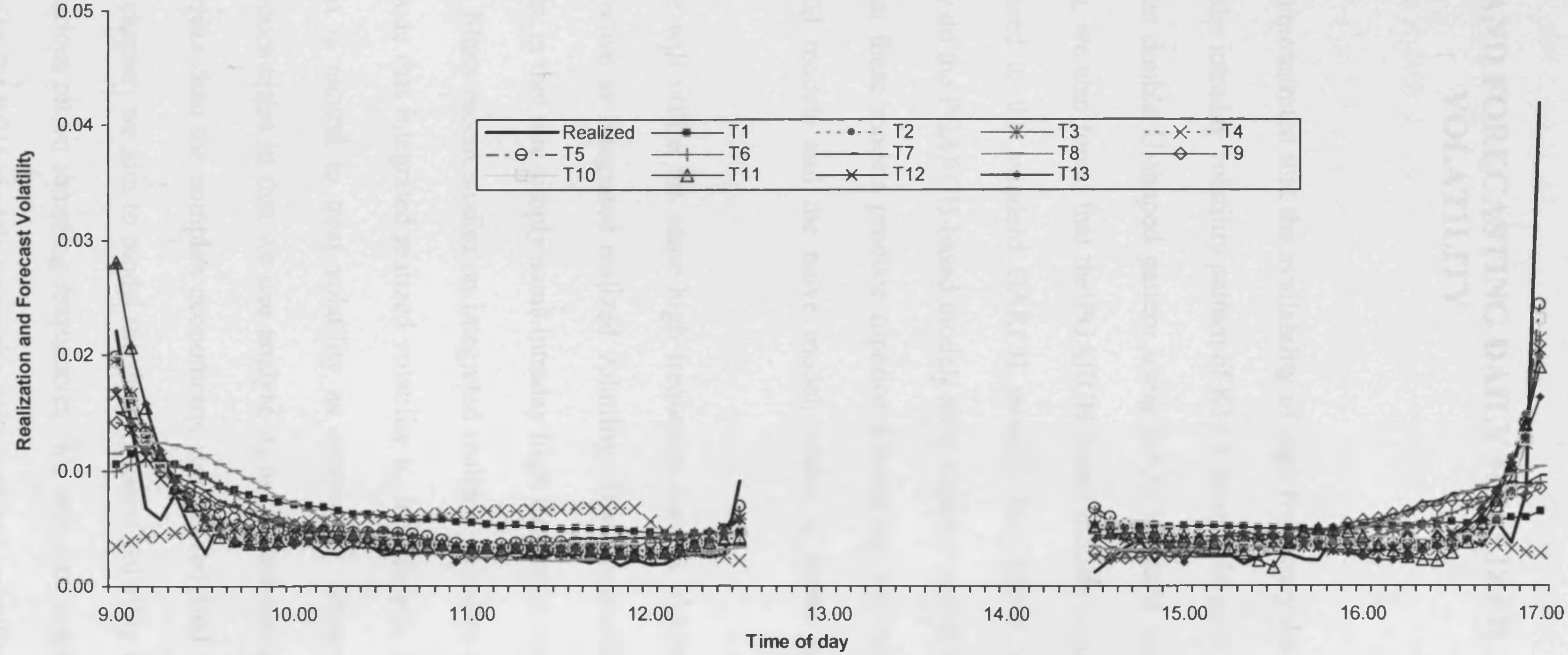


Figure 5.14: Forecasting Performance of Volatility Modelling Approaches

The chart below compares the plots of the mean realized volatility and the forecasted return volatilities (the daily average of 5-minute one-step-ahead forecasts) for a 100-Day out-of-sample forecasting period using the following approach:

T1 = Non-periodic GARCH model
T2 = Jointly estimated full dummy version of the PGARCH model
T3 = Two-step full dummy version of the PGARCH model
T4 = Jointly estimated partial dummy version of the PGRACH model
T5 = Two-step partial dummy version of the PGARCH model
T6 = Jointly estimated FFF version of the PGARCH model
T7 = Two-step FFF version of the PGARCH model
T8 = Jointly estimated Augmented FFF version of the PGARCH model
T9 = Two-step Augmented FFF version of the PGARCH model
T10 = Jointly estimated Spline version of the PGARCH model
T11 = Two-step Spline version of the PGARCH model
T12 = Jointly estimated Augmented Spline version of the PGARCH model
T13 = Two-step Augmented Spline version of the PGARCH model



CHAPTER 6

MODELLING AND FORECASTING DAILY REALIZED VOLATILITY

6.0 Introduction

In Chapter 4, we demonstrated that the availability of high frequency data has made it possible to model the intraday volatility pattern of KLCI returns. In particular, we were able to model the double U-shaped pattern using GARCH-based models. Using high frequency data, we also found that the PGARCH-based models produced superior model fit compared to the standard GARCH models. In Chapter 5, we demonstrated that not only do the PGARCH-based models have superior model fit in-sample; we also found that these models produce superior forecasting performance than the standard GARCH models and the naive model, which is based on the historical variance.

In this chapter, we will utilise the same high frequency data to construct a daily volatility measure known as integrated realized volatility. The construction of realized volatility is simple, in that one simply sums intraday high frequency squared returns, taken day by day. Many recent studies on integrated realized volatility using high frequency data conclude that integrated realized volatility is, in principle, error-free, and that therefore it is natural to treat volatility as observable. Observable volatility presents new opportunities in that we can analyse it, use it and forecast it with much simpler techniques than the complex econometric models required when volatility is latent. In this chapter, we aim to model the daily realized volatility using ARMA model based on various return sampling frequencies. We will also assess the performance of various daily GARCH models using the daily realized volatilities as the proxies for the true daily volatility. In addition, we will also evaluate whether both the ARMA and the GARCH models are able to produce accurate VaR measures.

This chapter is organized as follows. Section 6.1 gives an overview of the background of the recent research in realized volatility. In section 6.2, we provide a brief theoretical background of the quadratic variation theory that forms the basis for constructing the realized volatility. Section 6.3 describes the data and the methods for obtaining competing ARMA and GARCH forecasts. It also describes in detail the construction of the VaR models from the available ARMA and GARCH forecasts, as well as from the *RiskMetrics* model. In section 6.4, we present the estimation results and discuss the in-sample parameter estimates and out-of-sample forecast performance of alternative volatility models using various measures of realized daily volatility. In this section, we also assess the adequacy of VaR forecasts generated by the best performing ARMA and GARCH models at both the 1% and the 5% level of significance. Section 6.5 concludes the chapter. All results are reported at the end the chapter.

6.1 Chapter Background

It is a standard approach that the forecast performance of any volatility model is evaluated by comparing its predictions with realizations. Since volatility is not a directly observable process, this approach is not immediately applicable. The task of forecasting volatility is therefore difficult because of the need to identify the "true volatility" process. Identifying a suitable proxy for the true volatility is not an easy task but it is crucial. This is because any measure of volatility that represents the "true volatility" is used as the realized volatility against which the forecast performances of the volatility models are measured and subsequently evaluated. Lazarov (2004) argues that studies that employ an *ex post* estimate of volatility could induce a serious bias, because it favours the model which is used to calculate the estimate of latent volatility. Andersen, Bollerslev, Christoffersen and Diebold (2005) highlight the problem of finding the true volatility as follows:

> "Treating the volatility process as latent effectively transforms the volatility estimation problem into a filtering problem in which the "true" volatility cannot be determined exactly, but only extracted with some degree of error."
>
> (Andersen, Bollerslev, Christoffersen and Diebold, 2005, page 3)

Many researchers until recently have resorted to using daily squared returns, calculated from market daily closing prices, to proxy the true daily volatility. This is a non-parametric approach that is simple to compute[1] and widely used in the volatility forecasting literature. In Chapter 5, we have, in fact, used the squared returns as the proxy for true volatility in measuring the forecasting performances of the thirteen volatility-modelling approaches introduced in Chapter 4.

Prior to the availability of high-frequency data, the type of data that was most frequently used in association with the GARCH models was the daily market closing prices, from which daily returns are computed. It is therefore natural that the daily squared returns were often used as the proxy for true daily volatility. Using daily squared returns as the basis of forecast measurement, many of the earlier empirical results show that the parameters of different GARCH models are highly significant in-sample. However, the evidence is mixed regarding the provision of good out-of-sample forecasts. In fact, many of the more recent studies have shown that the standard GARCH models are incapable of producing good forecasts. Research findings by Jorion (1995, 1996), Figlewski (1997) and Andersen and Bollerslev (1998a), for example, show that the standard GARCH models provide poor forecasts

[1] Please refer to equation 5.1 for the computation of squared returns.

and explain little of the variability of the *ex post* daily squared returns measure. This naturally leads to the belief that the standard GARCH models may be of limited use in practice.

Andersen and Bollerslev (1998a), however, argue that the failure of the GARCH models to provide good forecasts is not a failure of the GARCH models *per se*, but a failure to specify correctly the true volatility measure against which forecasting performance is measured. It is argued that the standard approach of using the daily squared returns as the measure of the true volatility for daily forecasts is inappropriate because this measure includes a large and noisy independent zero mean constant variance error term, which is unrelated to actual volatility. Therefore, the daily squared returns measure is not a suitable estimator for the daily volatility and consequently does not provide a reliable estimate for the true underlying latent volatility. It is more likely to be for this reason that standard GARCH models often report poor predictive power.[2]

As an alternative, Andersen and Bollerslev (1998a) introduced a new generation of conditional volatility models, which make use of a volatility measure known as the integrated realized volatility. Use of such a measure allows more meaningful and accurate volatility forecast evaluation. The daily realized volatility can be constructed by summing up intraday squared returns. This allows the treatment of the daily volatility as observed rather than latent, providing that the sampling of high frequency squared returns is sufficiently frequent. By making use of the theory of quadratic variation and arbitrage-free processes, Andersen, Bollerslev, Diebold and Labys (2001, 2003) show that the realized volatility constructed as above is not only

[2] However, we have demonstrated in Chapter 5 that by using high frequency data, the performance of the standard GARCH models can be improved with the application of the PGARCH-based models.

model-free, but as the sampling frequency of the returns approaches infinity, the estimates are measurement-error-free as well.

Based on a simulation of realized volatility implied by the GARCH (1,1) diffusion limit, Andersen and Bollerslev (1998a) find that realized volatility provides a less noisy estimate of the latent volatility than does the daily squared returns. It is concluded that by sampling more frequently and producing a measure based on intraday data, the noisy component of the realized volatility diminishes, and that in theory, the realized volatility based on the high-frequency data is much closer to the actual volatility of the day. Subsequent studies by Barndorff-Nielsen and Shephard (2002a, 2002b) and Areal and Taylor (2002) indicate that the sum of squared high frequency intraday returns provides reliable estimation of the actual daily volatility.

There is also compelling evidence that volatility models that are parameterised using realized volatility produce superior forecasting performance. For example, Andersen, Bollerslev, Diebold and Labys (2003) consider the volatility of the Japanese Yen against the US Dollar and the Deutsche Mark against the US Dollar exchange rates, using an autoregressive fractionally integrated moving average (ARFIMA) model to characterize the realized volatility process. The results indicate that the predictive ability of this model is much better than the predictive ability of the GARCH (1,1) model, which relies on daily returns to compute the *ex post* estimate of the volatility. Lazarov (2004) estimates and compares several classes of volatility models for the DAX index futures, either using the realized variance or the squared daily returns. The findings show that realized variance is a much better estimate of the latent volatility than the sum of weighted daily squared returns and as such it is better suited for comparing the out-of-sample performance of competing volatility models. A similar conclusion is drawn when Bali and Lu (2005) apply the ARMA-fitted realized volatility models for 1-day-ahead and 20-day-ahead forecasts of the S&P 100 index. The results indicate that almost all information is provided for by the sum of squared five-minute returns. They conclude that there is little incremental information in the traditional volatility estimator based on the absolute demeaned daily index returns compared to those provided by the realized volatility measures. Studies on implied volatility have also highlighted the favourable results obtained when the realized volatility is used in time series volatility models. Results from the studies of Pong, Shackleton, Taylor and Xu (2002) and Lazarov (2004), for example, indicate that the realized volatility is a much more efficient estimator of the latent volatility than the daily returns, which enter as parameters in popular volatility models like the daily GARCH model and its various derivatives. It is not surprising, therefore, to find that more and more recent work on daily volatility modelling and forecasting has employed the realized volatility as a benchmark to which the volatility models' performances are compared and evaluated.

The main purpose of this chapter is to highlight the impact of using different *ex post* realized daily volatility measures (to proxy the true daily volatility) on the forecasting performance of competing volatility models, and the adequacy of the VaR models constructed from the available forecasts. Specifically, we will compare and assess the out-of-sample forecasting performances of two competing sets of volatility models. The first set of volatility models comprise the various GARCH models specified in the previous chapter. The GARCH models are estimated using the daily returns computed from the daily closing price at the end of each trading day. The GARCH forecasts are our primary focus because we want to see whether by using the daily realized volatility to proxy the actual daily volatility, one can obtain a better forecast performance over the forecasts measured against the traditional volatility proxy, i.e., the daily squared returns. The second set of volatility model is the ARMA (1,1) model, which is used to model the various daily realized volatility measures. The daily realized volatility is computed as the sum of the squared intraday returns for the given trading day. The motivation for this comparison arises from the desire to know whether by utilising intraday data (upon which the realized volatility is estimated), one can obtain a better model to proxy the true daily volatility, i.e., all the relevant data during the trading day are being compounded and accounted for. This is achieved by evaluating the performance of both the ARMA and GARCH models over a number of specially constructed daily realized volatilities. The ARMA model's forecasting performances in this instance could be useful in ascertaining the optimal intraday return sampling frequency for the daily realized volatility to be applied in the Malaysian market. We measure and evaluate the performance and the quality of the out-of-sample forecasts produced by the available volatility models, using both the MSFE and MAFE statistics and the Diebold and Mariano (1995) asymptotic test, respectively. Based on the forecasts obtained from the ARMA and GARCH volatility models, we construct the appropriate daily VaR models. We then assess the adequacy and quality of these daily VaR models at both the 1% and 5% level of significance.

This chapter complements the literature in two ways. First, we use high frequency data from an important emerging capital market, the KLSE, which is considered one of the biggest in South-East Asia. Second, we believe that this is the first study of its kind on the Malaysian stock exchange using ARMA and GARCH models to estimate and compare the properties of the realized daily volatilities.

6.2 An Overview of the Theoretical Background of Integrated Realized Volatility

A rigorous treatment of the theoretical background of this theory can be found in Andersen, Bollerslev, Diebold and Labys (2001, 2003). Let us consider the following simple multivariate continuous-time stochastic volatility diffusion process,

$$dp_t = \mu_t dt + Q d_t W_t . \tag{6.1}$$

where p_t is the $k \times 1$ instantaneous logarithmic price, μ_t, is a drift parameter, and $d_t W_t$ is a $k \times 1$ standard Brownian motion. The $k \times k$ positive definite diffusion matrix Q_t follows a strictly stationary process and satisfies $Q_t Q_t^{'} = \Omega_t$. For this diffusion, the integral of the instantaneous variances over the day, that is,

$$\tilde{\Omega}_t = \int^{t+1} \Omega_\omega d\omega . \tag{6.2}$$

provides an *ex post* measure of the true latent volatility associated with day *t*. By cumulating the intraday squared returns, as shown in Merton (1980), we can approximate the integrated volatility in equation (6.2) to any arbitrary precision. In particular, we can obtain an estimate, denoted by $\hat{\Omega}_t$, of $\tilde{\Omega}_t$ as

$$\hat{\Omega}_t = \sum_{j=1}^{\delta} r_{t+j/\delta} \cdot r_{t+j/\delta}^{'} \cdot \tag{6.3}$$

where $r_{t+j/\delta} \equiv p_{t+j/\delta} - p_{t+(j-1)/\delta}$ denotes the continuously compounded returns, sampled δ times per day. Note that the subscript *t* indexes the day, while *j* indexes the time within day *t*. The measure $\hat{\Omega}_t$ is referred to as realized volatility, as in Andersen, Bollerslev, Diebold and Labys (2001, 2003). By the theory of quadratic variation, it can be shown that equation (6.3) provides a consistent estimate of latent volatility as

$$\text{plim}_{\delta \to \infty} \hat{\Omega}_t = \tilde{\Omega}_t . \tag{6.4}$$

In other words, as the sampling frequency of returns increases, $\delta \to \infty$, the *ex post* realized volatility measures so constructed will converge to the integrated latent volatilities. This measure contrasts sharply with the common use of the squared *j*-period returns as the simple *ex post* volatility measure, which, although it provides an unbiased estimate for the realized volatility, is an extremely noisy estimator. Furthermore, for longer horizons, any conditional mean dependence will tend to contaminate this variance measure, whereas the mean component is irrelevant for the quadratic variation.

6.3 Data and Methodology

It is important to highlight in this chapter that there are two types of volatilities being examined. The first type of volatility is the integrated realized daily volatility, which is obtained by summing the intraday squared returns using the KLCI data. The daily integrated realized volatility, as explained above, is a volatility measure that is assumed to be model-free and an unbiased estimator of the true daily volatility. The choice of the appropriate frequency of intraday squared returns sampling is discussed below. The second type of volatility under consideration is the traditional measure of volatility based on the daily squared returns using the same set of data. This volatility measure is the most frequently used in the literature as a proxy for the true daily volatility.

The main focus of this chapter, therefore, is on the differences between the two measures, and whether there is any significant difference in the forecasting performance of volatility models that utilise these two different measures of volatility. For the purposes of meaningful comparisons and easier references, we also refer to the daily squared returns as the one-day frequency realized volatility. The ARMA

(1,1), GARCH (1,1), TGARCH (1,1) and the EGARCH (1,1) models are used extensively in this study. For easier references, we simply denote these models as the ARMA, GARCH, TGARCH and the EGARCH models respectively.

We set the stage by first computing various daily sums of intraday squared returns that will generate the daily realized volatilities, and compute the one-day frequency realized volatility (i.e. the daily squared returns) based on end-of-day returns. Next, we model all the realized daily volatilities using the ARMA model. We then generate one-day one-step-ahead forecasts from this model. Next, we employ the various GARCH models described above to model the daily volatility. In order to do this, we make use of the daily composite index end-of-day returns data as the input to the estimation process. Similarly, we generate one-day one-step-ahead forecasts using the in-sample parameters of the GARCH models. The finer details are explained below.

6.3.1 Modelling and Forecasting Realized Volatility

A point that has yet to be agreed upon in the construction of integrated realized volatility is the optimal frequency of intraday squared returns sampling, δ in equation (6.3) above. Earlier works such as French, Schwert and Stambaugh (1987) and Schwert (1989) obtained the monthly realized volatilities using daily return observations. With the arrival of high frequency data, many recent studies have experimented with different returns intervals of sizes ranging from one minute to 25 minutes. One of the earliest studies to use high frequency data is by Schwert (1990), who relied on the 15-minute returns to obtain the daily realized volatilities. However, several studies suggest that the choice of the optimal sampling frequency very much depends on the type of market being tested, market activity and the microstructure frictions associated with a particular market. It is important, therefore, that the sampling frequency considers a balance between measuring the volatility with as little noise as possible on one hand and avoiding market microstructure effects on the other. The microstructure effects are market frictions that arise due to market factors such as bid-ask price bounces, price discreteness or non-synchronous trading. As a trade-off between these two biases, Andersen, Bollerslev, Diebold and Labys (2001, 2003), for example, propose the use of 5-minute returns as the optimal sampling frequency in the US foreign exchange market.[3] Oomen (2001), on the other hand, argues that the optimal sampling frequency for his dataset (using FTSE-100 stock market index) is 25 minutes after evaluating the adequacy of sampling frequencies between 1 and 45 minutes. Giot and Laurent (2004), meanwhile, concur with Schwert (1990) and find that 15-minute returns are adequate for their work on data from the French CAC40 stock index and SP500 futures contracts traded on the Chicago Mercantile Exchange. Melvin and Melvin (2003) also use 15-minute returns to study the volatility spillovers of the Japanese Yen/US Dollar and the Deutsche Mark/US Dollar exchange rate across American, European and Asian markets.

In this study, we use 1-minute, 5-minute, 10-minute, 15-minute and 30-minute returns as the sampling frequencies in constructing the daily realized volatility estimates. We choose this range of return intervals so as to ascertain which return sampling frequency is the most appropriate for the Malaysian market. We also wish to examine the robustness of the 5-minute return sampling frequency in mitigating the problem of bias as suggested by Andersen, Bollerslev, Diebold and Labys (2001, 2003). Moreover, there is evidence that non-synchronous trading induces serial

[3] Similar suggestions are also advanced by Andersen, Bollerslev, Diebold and Ebens (2001) in the study of equity markets (New York Stock Exchange (NYSE), the American Stock Exchange (AMEX), and the National Association of Security Dealers Automated Quotation System (NASDAQ)) in the US.

correlation in the return process in many emerging markets such as the KLSE.[4] This in turn would render the cumulative squared returns measures as biased. Ebens (1999), however, argues that the microstructure effects are minimal when the focus is on an index, as the effects would tend to wash out in the aggregate. We will investigate whether this claim has any justification in the context of an emerging capital market.

The data used in this chapter are the KLCI in the form of 1-minute, 5-minute, 10-minute, 15-minute and 30-minute returns, as well as the end-of day returns, which will be used as the input in the GARCH model estimations and the computation of the one-day frequency realized volatility (i.e. the daily squared returns). Our sample covers the period from 29 January 2001 to 29 December 2002, resulting in a total of 406 trading days. In a typical trading day, the market opens at 9:00 am with a break for lunch at 12:30 pm. It then continues after lunch from 2:30 pm right through to 5:00 pm when the market closes for the day. This six-hour trading period provides us with a total of five sets of continuously compounded intraday returns for each day. The first set comprises of 360 continuously compounded 1-minute returns for each day, corresponding to $\delta = 360$ in the notation above. The second set makes use of 72 continuously compounded 5-minute returns for each day, corresponding to $\delta = 72$. The third set comprises of 36 continuously compounded 10-minute returns for each day, corresponding to $\delta = 36$. The fourth set is based on 24 continuously compounded 15-minute returns for each day, corresponding to $\delta = 24$. The fifth and final set is based on 12 continuously compounded 30-minute returns for each day, corresponding to $\delta = 12$. Based on the five returns series (obtained from the

[4] See, for example, Ariff, Shamsher and Annuar (1998).

logarithmic composite index difference), we construct the five competing daily realized volatilities (y_t) for the KLCI returns as

$$y_t = \sum_{j=1}^{\delta} r_{t+j/\delta}^2 \cdot \tag{6.5}$$

where r, t and j are defined as per equation (6.3) above and δ = 360, 72, 36, 24 and 12. We then examine the distributional characteristics of the daily realized volatilities for the sample period with reference to the mean, median, standard deviation, skewness, kurtosis and the normality of the distribution.

Next, the daily realized volatilities series are split into two sub-periods: an in-sample estimation period and an out-of-sample forecast evaluation period. The in-sample period covers the first 306 trading days, while the out-of-sample period comprises the last 100 trading days of the 406 trading day sample period.

6.3.2 Modelling and Forecasting Volatility using the ARMA model

As mentioned above, we model the various realized daily volatilities and one-day frequency realized volatility (i.e. the daily squared returns) using the ARMA (1,1) model (henceforth, we refer this as the ARMA model). The ARMA model, in this instance, postulates that the current value of the daily realized volatility series obtained from equation (6.5) depends linearly on its own previous value plus a combination of current and previous value of a white noise error term. The model could be written as follows:

$$y_t = \alpha + \beta_1 y_{t-1} + \varphi_1 u_{t-1} + u_t. \tag{6.6}$$

where u_t is a sequence of independently and identically distributed (i.i.d) random variables with $\mathrm{E}(u_t) = 0$, $\mathrm{E}(u_t^2) = \sigma^2$, and $\mathrm{E}(u_t u_s)$, $t \neq s$.

In the in-sample period, we fit the ARMA model to each of the five daily realized volatilities series and the one-day frequency realized volatility (i.e. the daily squared returns). The two information criteria we use in order to decide the appropriate model fit are the AIC and the SIC statistics. We will use this information to ascertain whether the best fitted model will produce superior out-of-sample forecasting performance later in the analysis.

In the out-of-sample period, based on the ARMA model, we generate one-day one-step-ahead forecasts for each of the five sets of the daily realized volatilities and the one-day frequency realized volatility series. To evaluate the performance and accuracy of the available forecasts, we apply the two forecast error statistics used in the previous chapter, namely MSFE and MAFE statistics. In order to get the appropriate error statistics, we compare the ARMA forecasts obtained against the particular daily realized volatility, which is used to proxy the true daily volatility. For example, the ARMA forecast obtained from the 1-minute return sampling frequency are compared against the corresponding 1-minute frequency daily realized volatility to get the appropriate MSFE and MAFE measures. Similarly, we apply this procedure for the other sampling frequencies (the 5-minute, the 10-minute, the 15-minute and the 30-minute return sampling frequencies) as well as for the one-day frequency realized volatility series.

6.3.3 Modelling and Forecasting Volatility using the GARCH models

In the estimation period, we make use of the various GARCH-based model specifications described earlier. All the GARCH-based models are estimated by maximum likelihood with Bollerslev-Wooldridge robust QML covariance/standard errors. Again we employ the AIC and the SIC statistics to evaluate the appropriate model fit. We then choose the GARCH-based model that gives the best model fit, based on the two information criteria, in order to ascertain whether the same model will produce superior forecasting results when we subsequently evaluate the forecasting performance of each model.

For the out-of-sample period, we generate one-day one-step-ahead forecasts for each of the 3 competing GARCH-based models. We then apply the two forecast error statistics used earlier: the MSFE and the MAFE. Similar to the approach adopted for the ARMA model, we compare the available GARCH forecasts against the five measures of the daily realized volatilities and the one-day frequency realized volatility series to obtain the appropriate MSFE and MAFE statistics. We begin by comparing the 3 competing GARCH forecasts against the 1-minute return frequency daily realized volatility used as the benchmark volatility to get the first set of MSFE and MAFE statistics. Next, using the same 3 GARCH forecasts, we compare these forecasts against the 5-minute return frequency daily realized volatility series to get the next sets of MSFE and MAFE statistics. We repeat this procedure with the other daily realized volatilities by comparing the same GARCH forecasts with the 10-minute return, 15-minute return, and 30-minute return sampling frequency daily realized volatility series to get additional sets of MSFE and MAFE measures. Finally, we compare the same GARCH forecasts against the one-day frequency realized volatility (i.e. the daily squared returns) to get the final set of MSFE and MAFE statistics. For each procedure, we select the best performing GARCH-based model by evaluating the model that produces the smallest MSFE and MAFE statistics respectively.

For comparison purposes, we also generate 5-day and 20-day one-step-ahead forecasts using 5-day and 20-day KLCI returns data described earlier using procedures described for the daily forecasting exercise above. To facilitate the estimation process, we only consider a sample of 400 days[5] instead of the original 406 days for the 5-day and 20-day forecasts. This provides us with 60 observations for the in-sample estimates and 20 observations for the out-of-sample forecasts for the 5-day forecasts and 15 observations for the in-sample estimates and 5 observations for the out-of-sample forecasts for the 20-day forecasts. We then compute the appropriate MSFE and MAFE for all available forecasts for according to the sampling frequencies discussed earlier.

6.3.4 Evaluating the Quality of Forecasts

We apply the Diebold and Mariano (1995) asymptotic test to test the null hypothesis of no difference in accuracy between the two competing forecasts. The properties of this test have been described in detail in the previous chapter. Since there are two types of volatility models being examined, we apply the tests separately for each type of daily realized volatility measure and the one-day frequency realized volatility (i.e. the daily squared returns). In the first undertaking, we focus on the 1-minute return sampling frequency daily realized volatility. We select the best performing GARCH-based model in terms of the MSFE statistics. Using the Diebold and Mariano (1995) test, we compare the forecast of this GARCH model with the forecast of the ARMA model. The null hypothesis is that the forecast generated by the ARMA model is of the same quality as the forecast generated by the GARCH model. The alternative hypothesis adopted is that the forecast produced by the ARMA model is superior to the forecast of the GARCH model. Next, for the same daily realized volatility, we now select the best GARCH-based model using the MAFE statistics and

[5] In this instance we omit the first three days and the last three days out of the 406 days original sample period to give us with a 400-day sample.

compare. We repeat the Diebold and Mariano (1995) test using the MAFE as the loss function by comparing the forecast of the best performing GARCH-based model with the forecast of the ARMA model.

The same procedure is then repeated for the best performing GARCH models in terms of producing the smallest MSFE and MAFE statistics with the ARMA forecasts for the next four daily realized volatilities, 5-minute, 10-minute, 15-minute and 30-minute return frequencies and finally for the one-day frequency realized volatility (i.e. the daily squared returns). The same objective is considered; that is, we test whether the forecast produced by the ARMA model is of the same quality as the forecast produced by the best performing GARCH model. For comparison purposes we also extend the Diebold and Mariano (1995) tests to all the volatility models for the 5-day and 20-day forecast evaluations.

6.3.5 Daily VaR Models

We follow the methodology described in the previous chapter to construct daily VaR forecasts at the 99% and 95% confidence levels. The VaR models are now assessed at the daily intervals instead of the 5-minute intervals used in the last chapter. In the first stage, we make use of the 3 one-day one-step-ahead GARCH volatility forecasts to produce 3 competing daily VaR forecasts. In order to evaluate the quality and adequacy of these VaR measures, we apply the framework for interval forecast evaluation developed by Christoffersen (1998). In this framework, we perform a test for "correct unconditional coverage", a test for "independence", and a test for "correct conditional coverage" for each of the VaR forecasts. In order to test for correct unconditional coverage, we test the null hypothesis (5.9) and apply the likelihood ratio statistic as specified in equation (5.14) in the previous chapter. We test for VaR levels of $\alpha = 0.01$ and $\alpha = 0.05$. We also employ the regression-based tests of Clement and Taylor (2003) to test for both the independence and the correct conditional coverage properties. This is done by performing an OLS of the indicator function (I_t) series on its one-lag value. The regression equation is specified in equation (5.18). We then test the null hypotheses (5.17) and (5.19) to test the quality and the adequacy of the VaR forecasts. Please refer to the previous chapter to appreciate the significance of each test and the hypotheses proposed to evaluate the adequacy of each VaR forecast.

In the second stage, we look at VaR models constructed from the ARMA forecast from each of the daily realized volatilities (the 1-minute, 5-minute, 10-minute, 15-minute and 30-minute return frequencies) and the one-day frequency realized volatility (i.e. the daily squared returns). In addition to the ARMA model, we also consider the *RiskMetrics* VaR model, the details of which have been discussed in Chapter 5. Similarly, in order to evaluate the quality and adequacy of these models, we apply the Christoffersen (1998) tests followed by the regression-based tests of Clement and Taylor (2003).

6.4 Results

6.4.1 Volatility Distribution Statistics

The summary statistics are presented in Table 6.1. The statistics report the results for the whole sample, which covers the period of 406 trading days. It can be observed that the mean of the daily squared returns is larger than the means of all the daily realized volatilities at 1.4339. Among the daily realized volatilities, the daily realized volatility based on the daily summation of 30-minute squared returns exhibits the highest mean value, at 1.209, while the lowest mean value of 0.9466 is produced by the daily realized volatility based on the 5-minute return sampling frequency. Turning to the median value, the daily realized volatility based on the 30-minute return sampling frequency again shows the highest value among all proxy volatilities at 0.5272. The one-day frequency realized volatility (i.e. the daily squared returns) exhibits the lowest median value at 0.3039. This series is also the most volatile series among all the proxy daily volatilities, with a maximum value of 39.8279 and a minimum value of zero. Among the daily realized volatilities, the largest maximum value is shown by the daily realized volatility based on the 5-minute return sampling frequency. However, the largest standard deviation value is exhibited by the daily realized volatility based on the 30-minute return sampling frequency.

The entire set of daily realized volatilities series is highly skewed. The series skewness coefficients for all the daily realized volatilities are positive, implying that the distributions of the volatilities are not symmetric but skewed to the right. The daily realized volatility based on the 1-minute return sampling frequency produces the largest skewness value, at 8.4296. The value of skewness for the daily squared returns is 5.8731, and this positions it at number four among the six daily realized volatilities. Looking at the coefficients of the series kurtosis, we find that all values are much larger than the normal value of 3, indicating that the distributions for all the daily volatilities series are highly leptokurtic. The largest kurtosis value is shown by the daily realized volatility based on the 1-minute return sampling frequency, at 102.1200. The value for the one-day frequency realized volatility (i.e. the daily squared returns) puts it in the fourth place among the six daily volatilities, at 44.6260. The Jarque-Bera normality test statistics for all daily realized volatilities are highly significant, with p-values of 0.0000 for all six series. This indicates that the null hypothesis of normality can be easily rejected for all the daily realized volatilities.

Figure 6.1 shows the plots of the five daily realized volatilities and the one-day frequency realized volatility (i.e. the daily squared returns) for the whole sample period of 406 days. It can be clearly observed that for all the proxy daily volatilities, the period of high volatility is approximately from days 60 to 90 of the sample period. The last 100 days are much less volatile than the first 100 days of the sample period. This is in line with the much more stable financial climate experienced by the Malaysian economy during the later period of the sample. It is also clear that the one-day frequency realized volatility (i.e. the daily squared returns) series is the most volatile among the entire set of daily realized volatilities studied here.

6.4.2 Model Fit

The model fit for the ARMA model for all the daily realized volatilities and the model fit for the daily GARCH models are presented in Tables 6.2. For the GARCH-based models, a particular volatility model is judged to be the best if the model fit produces the smallest AIC and SIC values. The GARCH model provides the best model fit with values of AIC of 3.2122 and SIC of 3.2609. This is in contrast to the findings in Chapter 4 and Chapter 5, where the EGARCH models clearly dominate both in the in-sample estimations and out-of-sample forecasting exercises.

6.4.3 Forecast Performance and Forecast Quality

The out-of-sample forecasting period covers a horizon of 100 days. The results are reported according to the particular daily volatility series used, i.e. the five different daily realized volatilities and the one-day frequency realized volatility (i.e. the daily squared returns).

Table 6.3 reports the MSFE and the MAFE statistics for the ARMA model and the GARCH models, respectively while Table 6.4 presents the MSFE and the MAFE for the ARMA model and the best performing GARCH-based model for each sampling frequency. For the 1-minute return sampling frequency, the MSFE and the MAFE statistics for the ARMA model are at 0.1119 and 0.3084 respectively. Among the GARCH-based models, we find that the EGARCH model produces the smallest MSFE statistics, at 1.0629 and the smallest MAFE statistics, at 0.8789. It is clear that the ARMA model perform better than the EGARCH model. This is not a surprise, considering that the GARCH estimates are based on the end-of-day returns and therefore may not be able to capture sufficiently the latent properties of the daily volatility, which in this case is represented by the 1-minute return sampling daily realized volatility. It is also interesting to note that the naive model performance is better than all the GARCH models' performances. This is true for both the MSFE and the MAFE statistics.

Similarly, for daily realized volatility based on the 5-minute return sampling frequency, as the benchmark volatility, the MSFE and the MAFE statistics for the ARMA model are at 0.1949 and 0.4195 respectively. For the GARCH-based models, the EGARCH model again produces the smallest MAFE statistics at 1.0961 and the smallest MAFE statistics at 0.8997. Again we see that the naive model forecasts outperform all the GARCH forecasts.

For daily realized volatility based on the 10-minute return sampling frequency the ARMA model reports MSFE and MAFE values of 0.2219 and 0.4338 respectively. There is no surprise for the performance of the GARCH-based models. As before, the EGARCH model outperforms the rest of the available GARCH-based models. The MSFE and the MAFE statistics for this model are at 1.0576 and 0.8907 respectively. It is interesting to observe that under the MAFE metric, we find that all the GARCH-based models actually outperform the naive model.

For forecasts using benchmark volatility based on the 15-minute return sampling frequency, the ARMA model produces MSFE metric with a forecast error of 0.2512 and MAFE metric with a forecast error of 0.4391. Among the GARCH-based models, the best performer for both the MSFE and the MAFE metric is the EGARCH model with values of 1.0242 and 0.8738 respectively. As before, we find that under the MAFE metric, the GARCH-based models easily outperform the naïve model.

For daily realized volatility based on the 30-minute return sampling frequency as the benchmark volatility the ARMA model reports MSFE metric with a value of 0.3529 and an MAFE value of 0.5237. Turning to the GARCH forecasts, the best performers are again the EGARCH model with an MSFE value of 0.9580 and an MAFE value of 0.8341. This time around, there is no question that all the GARCH forecasts outperform the unconditional variance forecasts. This is true for both the MSFE and MAFE statistics.

For the one-day frequency realized volatility the ARMA model produces an MSFE value of 2.2907 and an MAFE value of 1.1831 For the GARCH forecasts, the EGARCH model again outperforms the rest of the available GARCH-based models with an MSFE value of 2.3336 and an MAFE value of 1.0961. The results are consistent with the findings of Chapter 4 and Chapter 5. As expected, we find that all the GARCH forecasts outperform the unconditional variance forecasts for both the MSFE and MAFE metrics.

The plots of forecasts of the ARMA and the best performing GARCH-based models against each of the daily realized volatilities and one-day frequency realized volatility (i.e. the daily squared returns) are presented in Figures 6.2, 6.3, 6.4, 6.5, 6.6 and 6.7 respectively.

There are five observations that we would like to highlight here. The first observation is that the models with the best SIC value (as a result of the in-sample model estimation above) does not, in any case, provide superior forecasting performance as one might expect. For example, for the daily GARCH-based models, the in-sample estimation results indicate that the GARCH has the potential to produce the best forecast performance, considering that it is the model with the smallest SIC value of 3.2609. The out-of-sample forecast results instead find that the EGARCH model with an inferior SIC value of 3.2972, is able to produce superior forecasting performance.

The second observation we would like to highlight is regarding the performance of the ARMA models. We find that as the return sampling frequency becomes higher for the daily realized volatility, the forecasting performances of the ARMA models improve, i.e. the ARMA models produce smaller MSFE and MAFE statistics. For example, if we take the daily realized volatility based on the 30-minute return sampling frequency as the benchmark volatility, the MSFE and the MAFE figures for the ARMA model are 2.2907 and 1.1831, respectively. When we increase the return sampling frequency to 15 minutes, the MSFE and the MAFE for the ARMA model are now 0.2512 and 0.4391, respectively. The improvement in the forecasting performance of the ARMA model is clearly observable as we continue increasing the return sampling frequency. For example, for the 1-minute return sampling frequency, the MSFE and the MAFE statistics are 0.1119 and 0.3084 respectively. This observation is consistent with the theory of quadratic variation and arbitrage-free processes proposed by Andersen, Bollerslev, Christoffersen and

Diebold (2001, 2003).[6] It is also interesting to note that the ARMA model's forecasting performances are superior to the naive model (based on the mean realized volatility) for both MSFE and MAFE metrics, as well as for all measures of daily realized volatilities.

The third observation concerns the performance of the GARCH models. It can be observed that the forecasting performances of all the GARCH models are better when the measure of daily volatility is based on the summation of intraday squared returns, instead of the one-day frequency realized volatility (i.e. the daily squared returns). For example, if we refer to Table 6.3 and take the one-day frequency realized volatility as the benchmark daily volatility, the EGARCH model which provides the best forecasting performances produce MSFE and MAFE figures of 2.3336 and 1.0961, respectively. In contrast, if we apply the five daily realized volatilities as the benchmark volatilities, we find that the EGARCH models produce smaller figures for the MSFE and MAFE. The range of results for the MSFE and the MAFE are from 0.9580 to 1.0629 and from 0.8341 to 0.8789, respectively. In fact, the benchmark volatility that produces the best forecasting performance for the EGARCH model is the daily realized volatility with 30-minute return sampling frequency (MSFE figure of 0.9580 and MAFE figure of 0.8341).

The fourth observation regards the performance of the ARMA and the GARCH forecasts when the one-day frequency realized volatility (i.e. the daily squared returns) is used as the benchmark volatility to measure the MSFE and the MAFE statistics (see Tables 6.3 and 6.4 respectively). If we follow the performances of these two models based on the MSFE metric, it is clear that the ARMA model

[6] To recap, the theory predicts that as the sampling frequency of returns increases, the *ex post* realized volatility measures so constructed will converge to the integrated latent volatilities. Therefore, the improvement in the forecasting results of the ARMA models suggests that the estimate for the daily realized volatility is slowly converging to the true latent volatility.

outperforms all the GARCH-based models. This is clearly not a good result for the GARCH-based models, considering the fact that the ARMA model, with simpler structure, can outperform the supposedly superior GARCH formulations. However, the opposite is observed if we consider the MAFE metric. The GARCH-based models perform better on the whole when compared with the performances of the ARMA model.

The fifth and final observation is with regards to the most appropriate intraday squared return sampling frequency for the Malaysian market. In order to determine the optimal return sampling frequency, we plot the graphical diagnostic termed the "volatility signature plot" developed by Andersen, Bollerslev, Diebold and Labys (2001). This is a plot of average realized volatility against return sampling frequency, which may reveal the severity of microstructure[7] bias as sampling frequency increases, and may therefore be useful in guiding the selection of sampling frequency. In Figure 6.8, we show the plots of the average daily realized volatilities against the sampling frequencies used in this study. The average daily volatility is at its lowest when the sampling frequency is at 5 minutes. Therefore, we would recommend the use of a return sampling frequency of 5 minutes, which represents a reasonable trade-off between minimizing microstructural bias and minimizing sampling error. The result is consistent with the findings of Andersen, Bollerslev, Diebold and Labys (2001, 2003), who suggest the use of 5-minute return frequency for the US foreign exchange market.

Table 6.5 presents the results of the DM test for forecast quality. The DM tests performed using the MSFE and the MAFE as the loss functions provide significant results for all comparisons of the ARMA model and the corresponding EGARCH

[7] According to Andersen, Bollerslev, Diebold, and Labys (2001), the optimal sampling frequency will likely be a value ideally high enough to produce a volatility estimate with negligible sampling variation, yet low enough to avoid microstructure bias.

model for all daily realized volatilities based on the summation of intraday squared returns. The results suggest that the quality of forecasts produced by the ARMA model in the pair-wise comparisons is superior to the forecasts generated by the EGARCH model. In contrast, the pair-wise comparisons between the ARMA and the EGARCH model for the one-day frequency realized volatility (i.e. the daily squared returns) produce insignificant result for the MSFE metric, suggesting that the quality of forecasts for both the ARMA and the EGARCH models is the same. However, the results for the MAFE metric indicate that the quality of the EGARCH forecast is superior to the quality of the ARMA forecast at the 5% level.

Tables 6.6 and 6.8 summarise the forecasting performances of the ARMA and the GARCH-based models for the 5-day and the 20-day one-step-ahead out-of-sample forecasts respectively. For the 5-day forecasts, the best performing GARCH-based model is the TGARCH model, which dominates all the return sampling frequencies and similarly, for the 20-day forecasts, the GARCH model is clearly dominant. These results are not consistent with the one-day one-step-ahead forecasts in which the EGARCH model dominates all GARCH forecasts comparisons regardless of the choice of return sampling frequency. For the 5-day forecasts, we find that as the return sampling frequency becomes higher for the realized volatility, the forecasting performances of the ARMA model improve, i.e. the ARMA models produce smaller MSFE and MAFE statistics. This is consistent with the second observation for the one-day one-step-ahead forecasts above, There is somewhat mixed performance of the GARCH-based models with regards to the choice of return sampling frequencies for the 5-day forecasts. It can be observed that the forecasting performances of all the GARCH-based models are better when the measure of realized volatility is based on the 1-minute, 5-minute, 10-minute and 15-minute summation of intraday squared returns, instead of the 5-day frequency realized volatility (i.e. the 5-day squared returns). However, the same could not be said for the GARCH-based models using the 30-minute return sampling frequency measure. The results show that all the GARCH-based models performed worse than similar models using the 5-day frequency realized volatility measure. It is also clear that in most cases, the ARMA[8] and the naive models produce smaller forecast errors than the GARCH-based models. For the 20-day forecasts, the results are mostly consistent with the observations for the 5-day forecasts. We observe that in all comparisons, the ARMA and the naïve models perform better than the GARCH-based models in terms of producing smaller forecast errors. This is true for all return sampling frequencies. We also find that there is an improvement in the performances of the GARCH-based models that employ the intraday realized volatility measures over similar models that are based on the 20-day squared returns measure.

Tables 6.7 and 6.9 report the results of the DM tests for the 5-day and 20-day forecasts respectively. For the 5-day forecast, the DM tests performed using the MSFE as the loss function provide significant results at the 5% level for all comparisons between paired ARMA and GARCH-based models. This suggests that the quality of forecast between competing GARCH-based models is not the same i.e. of different quality. The same could not be said for the MAFE metric. In this instance, only comparisons using the 1-minute and 5-minute return sampling frequencies produce significant results at the 5% level, while for other return sampling frequencies, the results are insignificant, which suggest that the quality of forecasts for each pair is the same. For the 20-day forecasts, only comparisons using the 1-minute, 5-minute and the 20-day return sampling frequencies produced significant

[8] However, for the 5-day frequency realized volatility, the TGARCH model produces smaller MAFE figure than the ARMA model.

results at the 5% level. This is true for both the MSFE and MAFE statistics. Therefore, we conclude for that these three cases, the quality of the ARMA forecast is superior to the quality of the GARCH forecast.

6.4.4 VaR Performance

We proceed with the results of the tests for correct unconditional coverage and then discuss the results of the regression tests of Clements and Taylor (2003) for both the independence and correct conditional coverage tests. The actual daily returns for the out-of-sample period are used as the benchmarks to produce the indicator function I_t series described in the previous chapter.

6.4.4.1 Daily GARCH Models

Test for "Correct Unconditional Coverage" H0: $\hat{f} = \alpha$

Table 6.10 presents the results for the 99% VaR coverage $(\alpha = 0.01)$, while Table 6.13 reports the results for the 95% VaR coverage $((\alpha = 0.05)$ for the evaluation of the 3 VaR models constructed from the 3 available GARCH out-of-sample forecasts. From both tables, we find that in all cases, the likelihood ratio statistics obtained are statistically insignificant at the 1% level Therefore, we accept the null hypothesis (5.9) and reject the alternative hypothesis (5.10), i.e., the observed failure rate $(\hat{f})$ in all cases is the same as the required failure rate (α) as specified in the VaR model. We conclude that for all these GARCH-based models, the correct unconditional coverage for the 99% VaR and 95% VaR models is satisfied and appears to be adequate.

Test for "Independence"

The results for the 99% and 95% VaR independence tests for the VaR models are presented in Table 6.11 and Table 6.14, respectively. The regression tests for the 99% and 95% VaR coverages produce positive results. All the GARCH models appear to have the property of independence because the F-statistics are not statistically significant at either the 1% or 5% levels. Therefore, these models appear adequate for the 99% and the 95% VaR models and we accept the null hypothesis (5.19).

Test for "Correct Conditional Coverage"

The results for both the 99% and 95% VaR measures are presented in Table 6.12 and Table 6.15, respectively. The outcomes of the regression tests for both the 99% and the 95% VaR coverages mirror the results for the "independence" test above. All the GARCH-based models appear adequate. The overall results for the 99% VaR coverage indicate that for all models, the F-statistics are not statistically significant at either the 1% or 5% levels. Therefore, we accept hypothesis (5.17) and conclude that all the GARCH-based VaR models have the property of correct conditional coverage. The same conclusion can be drawn regarding the GARCH-based models for the 95% VaR coverage. For these models, the F-statistics are insignificant at the 5% level. The regression tests do not exhibit the existence of significant lagged effects in the failure process. Therefore, we have to accept the null hypothesis (5.17) and conclude that the models are adequate at providing the required 95% VaR coverage.

6.4.4.2 Daily ARMA and *RiskMetrics* Models – Daily Realized Volatility and Daily Squared Returns

Tables 6.16 to 6.18 present the results for the 99% VaR coverage $(\alpha = 0.01)$, while Tables 6.19 to 6.21 report the results for the 95% VaR coverage ($(\alpha = 0.05)$ for the evaluation of the VaR models constructed from the ARMA forecasts measured against the six daily realized volatilities.

The overall results indicate that the ARMA model has the appropriate unconditional coverage, independence and the correct conditional properties. This is true for both the 99% and the 95% VaR coverages. All coverage tests yield insignificant statistics at both the 1% and the 5% levels regardless of the return sampling frequency used. The results for the *RiskMetrics* model are quite similar though it fails the correct unconditional coverage at the 1% level. Therefore, we conclude that for both the ARMA and the *RiskMetrics* models, the VaR models appear to be adequate and accurate.

It is clear that the daily GARCH models employed in this chapter produce more accurate and reliable VaR models than the non-periodic GARCH models in Chapter 5, assuming that the distribution of returns series is normal. At both the 99% and 95% VaR coverage, the effects of fat tails are stronger for models that are based on high frequency data. Therefore, we see a rapid deterioration in performance not only for the non-periodic GARCH models, but also for the PGARCH-based and the *RiskMetrics* models. It appears that at the daily level, the effects of fat tails are insignificant, thus the strong performances of the daily GARCH, the *RiskMetrics*, and the ARMA models.

We do not provide any VaR analysis for the 5-day and 20-day forecasts in view of the very small[9] number of out-of-sample observations obtained from the

[9] For the 5-day forecasts, only 20 out-of-sample observations are obtained while for the 20-day forecast, only 5 out-of-sample observations are available. This is in relation to the 100-day forecasting period applied in the study.

modified sample (see section 6.3.3) above. We believe that the small number of out-of-sample observations[10] would provide inaccurate VaR models.

6.5 Conclusion

The choice of the *ex post* estimate of volatility is crucial in the tests performed in this chapter. In this chapter, we focus on two types of volatility to proxy the actual daily volatility. The first type of volatility being examined here is the daily realized volatility. This is computed as a series of daily sums of intraday squared returns: specifically, the 1-minute, 5-minute, 10-minute, 15-minute and 30-minute intraday squared returns are used to produce five competing daily realized volatilities. Andersen, Bollerslev, Diebold and Labys (2001, 2003) provide the theoretical foundation (the theory of quadratic variation and arbitrage-free processes) and justification for this measure of daily realized volatility. They show that this measure provides consistent and reliable estimates of the unobservable daily volatility. Many recent studies have demonstrated the efficiency of this measure. The main appeal of this daily volatility measure is that it incorporates the intraday volatility components, which are not considered and are missing in the daily squared returns computations.

The second type of volatility examined is the daily squared returns, also known as the one-day frequency realized volatility. This is the traditional method of measuring the daily volatility. There is, however, no sound theory to justify this method apart from it being a simple estimator of volatility. Consequently, it is very popular and has become the mainstay of many studies in volatility modelling and forecasting. However, Andersen and Bollerslev (1998a, 1998b) argue that this method is a noisy estimator for daily volatility and therefore it does not provide a reliable

[10] The Basle Committee (1998) recommends a backtest which sets the market risk capital requirements equal or greater than the average of the daily VaR measures during the preceding sixty business days, times the supervisory multiplier set by the committee.

estimate for the true underlying latent volatility. Andersen and Bollerslev (1998a) also suggest that the recent poor forecasting performance of the standard GARCH models is partly due to the use of the daily squared returns as the benchmark volatility to measure forecast errors.

The results obtained in this chapter demonstrate the superiority of the daily realized volatility measure over the daily squared returns measure. The one-day frequency realized volatility (i.e. the daily squared returns) series is evidently more volatile than the five daily realized volatility series. This can be observed from the summary statistics of the six volatility series in Table 6.1 and from the plots of the series in Figure 6.1. We find that the GARCH-based models produce superior forecasting performance when the benchmark volatilities used are the five daily realized volatilities (which are based on the summation of intraday squared returns), instead of the one-day frequency realized volatility (i.e. the daily squared returns). We would also recommend the 5-minute return as the optimal sampling frequency for the daily realized volatility among the five different return sampling frequencies examined for the Malaysian market here. This is in line with the optimal 5-minute return sampling frequency recommended by Andersen, Bollerslev, Diebold and Labys (2001). In view of the better forecasting performance produced by the daily GARCH models when the daily realized volatility based on the summation of intraday squared returns is used as the daily volatility measure, we support the findings that the daily squared returns is not a reliable *ex post* estimate of the true daily volatility, and therefore, wherever possible, it should be substituted with the realized volatility measure considered in this study.

The ARMA model, which is used to model the daily realized volatility, certainly produce superior forecasting performance compared to the various GARCH forecasts when all five daily realized volatilities are used as the benchmark volatilities to measure the forecast errors. However, the forecasting performance of the ARMA model is inferior to the forecasting performances of the GARCH-based models when we consider the MAFE metric for the one-day frequency realized volatility (i.e. the daily squared returns as the benchmark volatility). We also find that as the return sampling frequency of the daily realized volatility becomes higher (from one-day frequency to 30-minute to 15-minute to 10-minute etc.), the forecasting performance of the various ARMA models improves considerably and the size of the forecast errors produced also becomes smaller. This is consistent with the theory of quadratic variation and arbitrage-free processes discussed earlier.

The Diebold and Mariano (1995) tests applied to the forecasts suggest that the quality of the ARMA forecast is superior to the quality of the EGARCH forecast when the same daily realized volatility (based on summation of intraday squared returns) is used as the benchmark volatility. However, the opposite is true if we consider the one-day frequency realized volatility. The results suggest that both the ARMA and EGARCH forecasts are of the same quality when the MSFE is considered as the loss function. However, the situation is the opposite when the MAFE is used as the loss function. The results indicate that the quality of the EGARCH forecast is superior to the quality of the ARMA forecast.

The overall results for the 5-day and 20-day one-step-ahead forecasts are consistent with the one-day one-step-ahead forecasts. Generally, the ARMA produces superior forecasting performances compared to the performances of the GARCH-based models. The results also indicate the merit of using the intraday summation of squared returns in producing better performances from the GARCH-based models. However, the majority of the results of the differences in the quality of forecast between competing ARMA and GARCH-based models are not as significant as those obtained for the one-day one-step-ahead forecasts.

The VaR models constructed from the GARCH, the ARMA and the *RiskMetrics* forecasts appear to satisfy all the requirements of the framework for interval forecast evaluation at both the 99% and the 95% VaR coverage and, therefore, they are economically accurate and reliable. The overall results suggest that the daily GARCH-based, the *RiskMetrics*-based, and the ARMA-based VaR models investigated in this chapter are more accurate than the standard GARCH, PGARCH, and the *RiskMetrics* based VaR models evaluated in Chapter 5. This could be partly explained by the less severe fat-tail effects experienced at the daily level.

The work done in this chapter completes the final investigation towards a better understanding of the dynamics of intraday volatility on the KLSE. In this chapter, we have demonstrated the benefits of using the daily realized volatility measure as the *ex post* true daily volatility measure. It is simple to compute and could be modelled adequately using simple ARMA models. In addition, the realized volatility measures improve the forecasting performances of the standard daily GARCH models. The application of the realized volatility measures also produces accurate and adequate ARMA-based as well as GARCH-based VaR models. In the next chapter, we will summarise the major findings of our investigations into the dynamic characteristics of the intraday return volatility on the KLSE.

Table 6.1: Summary Statistics for the Realized Volatility and the Daily Squared Returns (For the whole sample – 406 Trading days) KLCI Data

This table reports the summary statistics for the daily realized volatilities based on the 1-minute, 5-minute, 10-minute, 15-minute and 30-minute returns frequencies of the sample under study. It also reports the summary statistics for the demeaned daily squared returns, which are computed based on the end-of-day prices for the same sample.

	Daily Realized Volatility Return Frequency					One-day frequency Realized Volatility/ Daily Squared Returns
	1 min	**5 min**	**10 min**	**15 min**	**30 min**	
Mean	0.9637	0.9466	1.0506	1.0847	1.2097	1.4339
Median	0.5054	0.4445	0.4633	0.4827	0.5272	0.3039
Maximum	25.2119	27.6839	27.0661	20.7921	22.1117	39.8279
Minimum	0.1163	0.0805	0.07076	0.0377	0.0298	0.0000
Standard Deviation	1.7443	1.9129	2.0976	2.0415	2.4226	3.8713
Skewness	8.4296	8.4086	6.8615	5.2855	5.3302	5.8731
Kurtosis	102.1200	102.2655	68.9919	38.3051	37.1067	44.6260
Jarque-Bera (*p*-value)	171011 (0.0000)	171475 (0.0000)	76857 (0.0000)	22976 (0.0000)	21601 (0.0000)	31646 (0.0000)

Table 6.2: Model Fit for Auto Regressive Moving Average (ARMA) models and Generalised Autoregressive Conditionally Heteroskedastic (GARCH) models KLCI Sum-of-Squared Returns (In-sample 306-day) Daily Realized Volatility

This table reports the log likelihood (LL), the Akaike Information Criterion (AIC), and the Schwarz Information Criterion (SIC) for the ARMA models that employ the sum of squared returns sampling frequency and the daily squared returns to estimate the daily realized volatility. This table also reports the LL, the AIC and the SIC for the GARCH models below. The conditional volatilities are estimated based on 306 end-of-day returns.

MODEL	LL	AIC	SIC
<u>**ARMA Models**</u>			
Daily Realized Volatility estimate based on:			
1-minute Returns Sampling Frequency			
Mean Realized Volatility	-640.4622	4.1926	4.2047
ARMA	-628.6604	4.1420	4.1786
5-minute Returns Sampling Frequency			
Mean Realized Volatility	-669.7809	4.3842	4.3964
ARMA	-655.6030	4.3187	4.3553
10-minute Returns Sampling Frequency			
Mean Realized Volatility	-697.4518	4.5650	4.5772
ARMA	-672.4658	4.4293	4.4659
15-minute Returns Sampling Frequency			
Mean Realized Volatility	-688.3604	4.5056	4.5178
ARMA	-654.1980	4.3095	4.3461
30-minute Returns Sampling Frequency			
Mean Realized Volatility	-741.4458	4.8526	4.8648
ARMA	-701.6744	4.6208	4.6574
One-day Frequency Realized Volatility			
Mean Realized Volatility	-884.5195	5.7877	5.7999
ARMA	-875.3312	5.7595	5.7961
<u>**Daily GARCH Models**</u>			
GARCH	-485.8579	3.2122	3.2609
TGARCH	-485.7423	3.2179	3.2789
EGARCH	-488.5177	3.2362	3.2972

Table 6.3: Forecast Performance 100-Day Daily Forecast Horizon – Auto Regressive Moving Average (ARMA) Models and Generalised Autoregressive Conditionally Heteroskedastic (GARCH) models

This table reports two forecast error statistics for the forecasts produced by the ARMA and the GARCH models below. The results are based on a one day one-step-ahead forecast covering a 100-day out-of-sample period. The errors computed are the mean squared forecast error (MSFE) and the mean absolute forecast error (MAFE). The benchmark daily volatility is represented by the appropriate daily realized volatility (based on the return sampling frequency). The GARCH forecasts are estimated from the in-sample model fit based on the end-of-day returns.

MODEL	MSFE	MAFE
Benchmark Daily Realized Volatility based on:		
1-minute Returns Sampling Frequency		
Naive Model	0.7039	0.8211
ARMA	0.1119	0.3084
GARCH	1.0745	0.8830
TGARCH	1.0694	0.8950
EGARCH	1.0629	0.8789
5-minute Returns Sampling Frequency		
Naive Model	0.7056	0.8144
ARMA	0.1949	0.4195
GARCH	1.1071	0.9005
TGARCH	1.1006	0.9116
EGARCH	1.0961	0.8997
10-minute Returns Sampling Frequency		
Naive Model	0.9201	0.9263
ARMA	0.2219	0.4338
GARCH	1.0695	0.8915
TGARCH	1.0617	0.9027
EGARCH	1.0576	0.8907
15-minute Returns Sampling Frequency		
Naive Model	0.9651	0.9451
ARMA	0.2512	0.4391
GARCH	1.0364	0.8746
TGARCH	1.0281	0.8858
EGARCH	1.0242	0.8738
30-minute Returns Sampling Frequency		
Naive Model	1.2113	1.0449
ARMA	0.3529	0.5237
GARCH	0.9772	0.8388
TGARCH	0.9642	0.8486
EGARCH	0.9580	0.8341
One-day Frequency Realized Volatility		
Naive Model	2.8547	1.4539
ARMA	2.2907	1.1831
GARCH	2.4138	1.1148
TGARCH	2.3815	1.1178
EGARCH	2.3336	1.0961

**Table 6.4: Best Forecast Performance – ARMA and GARCH Models
Daily Realized Volatility and Daily Squared Returns**

This table reports the models that produced the forecasting performance in terms of the mean squared forecast error (MSFE) and the mean absolute forecast error (MAFE) statistics. The results for the daily realized volatility are shown according to the return frequency used to estimate the daily volatility and which are subsequently used in the measurement of forecast errors. The best performing GARCH model is reported for each return frequency used as the benchmark volatility to measure the MSFE and the MAFE. The GARCH models are estimated based on end-of-day returns.

	MSFE	MAFE
Daily Realized Volatility based on:		
1-minute Return Frequency		
ARMA	0.1119	0.3084
EGARCH	1.0629	0.8789
5-minute Return Frequency		
ARMA	0.1949	0.4195
EGARCH	1.0961	0.8997
10-minute Return Frequency		
ARMA	0.2219	0.4338
EGARCH	1.0576	0.8907
15-minute Return Frequency		
ARMA	0.2512	0.4391
EGARCH	1.0242	0.8738
30-minute Return Frequency		
ARMA	0.3529	0.5237
EGARCH	0.9580	0.8341
One-day Return Frequency		
ARMA	2.2907	1.1831
EGARCH	2.3336	1.0961

Table 6.5: Comparing Forecast quality – 100-Day Daily Forecast Horizon
Daily Realized Volatility

This table reports the results of the Diebold and Mariano (1995) asymptotic test for forecast quality evaluation. The null hypothesis is that the forecasts generated by the ARMA models are of the same quality as the forecasts generated by the GARCH models. The alternative hypothesis adopted is that the forecasts produced by the ARMA models are superior to the forecasts produced by the GARCH models. The results are based on the comparison of the appropriate ARMA and GARCH models, as listed in Table 6.4. The test is implemented with the mean squared forecast error (MSFE) and the mean absolute forecast error (MAFE). The significance of these tests are denoted by ** (1% significance) and * (5% significance). The true daily volatility is proxied by the realized daily volatility.

Comparison	Metric	
	MSFE	MAFE
DAILY REALIZED VOLATILITY		
1-minute Return Frequency		
ARMA v. EGARCH	-5.4419**	-9.8909**
5-minute Return Frequency		
ARMA v. EGARCH	-5.0925**	-8.5849**
10-minute Return Frequency		
ARMA v. EGARCH	-5.0169**	-8.4756**
15-minute Return Frequency		
ARMA v. EGARCH	-4.9837**	-8.2388**
30-minute Return Frequency		
ARMA v. EGARCH	-4.0938**	-6.3583**
One-day Return Frequency		
ARMA v. EGARCH	-0.2974	2.4980*

Table 6.6: Forecast Performance 100-Day Forecast Horizon – Auto Regressive Moving Average (ARMA) Models and Generalised Autoregressive Conditionally Heteroskedastic (GARCH) models – 5-day One-step-ahead Out-of-sample Forecasts

This table reports two forecast error statistics for the forecasts produced by the ARMA and the GARCH models below. The results are based on a 5-day one-step-ahead forecast covering a 100-day out-of-sample period. The errors computed are the mean squared forecast error (MSFE) and the mean absolute forecast error (MAFE). The benchmark daily volatility is represented by the appropriate daily realized volatility (based on the return sampling frequency). The GARCH forecasts are estimated from the in-sample model fit based on the 5-day returns.

MODEL	MSFE	MAFE
Benchmark 5-Day Realized Volatility based on:		
1-minute Returns Sampling Frequency		
Naive Model	8.1885	0.8049
ARMA	0.4749	0.0516
GARCH	10.7454	0.7209
TGARCH	9.9302	0.6889
EGARCH	11.3489	0.7436
5-minute Returns Sampling Frequency		
Naive Model	5.6942	0.7197
ARMA	2.0824	0.4272
GARCH	13.6412	0.7782
TGARCH	12.7245	0.7463
EGARCH	14.3174	0.8010
10-minute Returns Sampling Frequency		
Naive Model	8.8322	0.9016
ARMA	3.4906	0.5587
GARCH	14.5040	0.7603
TGARCH	13.5639	0.7284
EGARCH	15.1969	0.7831
15-minute Returns Sampling Frequency		
Naive Model	9.6855	0.9713
ARMA	4.5571	0.6589
GARCH	15.1645	0.7339
TGARCH	14.2086	0.7021
EGARCH	15.8686	0.7567
30-minute Returns Sampling Frequency		
Naive Model	11.5273	1.0759
ARMA	5.8826	0.7564
GARCH	19.6592	0.8131
TGARCH	18.5749	0.7812
EGARCH	20.4553	0.8359
5-day Frequency Realized Volatility		
Naive Model	6.7229	0.8545
ARMA	6.7149	0.7731
GARCH	17.9446	0.7844
TGARCH	17.1979	0.7525
EGARCH	18.4991	0.8072

Table 6.7: Comparing Forecast quality – 100-Day Forecast Horizon
5-Day Realized Volatility

This table reports the results of the Diebold and Mariano (1995) asymptotic test for forecast quality evaluation. The null hypothesis is that the forecasts generated by the ARMA models are of the same quality as the forecasts generated by the GARCH models. The alternative hypothesis adopted is that the forecasts produced by the ARMA models are superior to the forecasts produced by the GARCH models. The results are based on the comparison of the appropriate ARMA and GARCH models, as listed in Table 6.6. The test is implemented with the mean squared forecast error (MSFE) and the mean absolute forecast error (MAFE). The significance of these tests are denoted by ** (1% significance) and * (5% significance). The true daily volatility is proxied by the realized daily volatility.

Comparison	Metric	
	MSFE	**MAFE**
5-DAY REALIZED VOLATILITY BASED ON		
1-minute Return Frequency		
ARMA v. TGARCH	-21.7767**	-2.8253**
5-minute Return Frequency		
ARMA v. TGARCH	-10.6926**	-2.0049*
10-minute Return Frequency		
ARMA v. TGARCH	-6.9649**	-1.2631
15-minute Return Frequency		
ARMA v. TGARCH	-5.2771**	-0.3570
30-minute Return Frequency		
ARMA v. TGARCH	-4.6949**	-0.1921
5-day Return Frequency		
ARMA v. TGARCH	-3.2974**	1.3334

Table 6.8: Forecast Performance 100-Day Forecast Horizon – Auto Regressive Moving Average (ARMA) Models and Generalised Autoregressive Conditionally Heteroskedastic (GARCH) models – 20-Day One-step-ahead Out-of-sample Forecasts

This table reports two forecast error statistics for the forecasts produced by the ARMA and the GARCH models below. The results are based on a 20-day one-step-ahead forecast covering a 100-day out-of-sample period. The errors computed are the mean squared forecast error (MSFE) and the mean absolute forecast error (MAFE). The benchmark daily volatility is represented by the appropriate daily realized volatility (based on the return sampling frequency). The GARCH forecasts are estimated from the in-sample model fit based on the 20-day returns.

MODEL	MSFE	MAFE
Benchmark 20-Day Realized Volatility based on:		
1-minute Returns Sampling Frequency		
Naive Model	138.4826	11.7236
ARMA	6.5765	1.7879
GARCH	99.4709	9.9212
TGARCH	106.7647	10.2822
EGARCH	155.0247	12.4090
5-minute Returns Sampling Frequency		
Naive Model	96.5298	9.7199
ARMA	63.2736	7.8442
GARCH	133.9547	11.4849
TGARCH	142.3774	11.8459
EGARCH	197.2886	13.9727
10-minute Returns Sampling Frequency		
Naive Model	147.7144	12.0044
ARMA	113.8831	10.5088
GARCH	143.1551	11.8129
TGARCH	151.8148	12.1739
EGARCH	208.1214	14.3008
15-minute Returns Sampling Frequency		
Naive Model	160.7909	12.4681
ARMA	132.5782	11.2853
GARCH	150.1923	12.0356
TGARCH	159.0127	12.3966
EGARCH	216.2665	14.5234
30-minute Returns Sampling Frequency		
Naive Model	188.4061	13.4350
ARMA	174.6798	13.1354
GARCH	204.0075	14.0036
TGARCH	214.2488	14.3646
EGARCH	279.8736	16.4914
20-day Frequency Realized Volatility		
Naive Model	287.9076	18.6590
ARMA	270.5780	15.8302
GARCH	546.5138	23.0994
TGARCH	563.3227	23.4604
EGARCH	667.6377	25.5873

Table 6.9: Comparing Forecast quality – 100-Day Forecast Horizon
20-Day Realized Volatility

This table reports the results of the Diebold and Mariano (1995) asymptotic test for forecast quality evaluation. The null hypothesis is that the forecasts generated by the ARMA models are of the same quality as the forecasts generated by the GARCH models. The alternative hypothesis adopted is that the forecasts produced by the ARMA models are superior to the forecasts produced by the GARCH models. The results are based on the comparison of the appropriate ARMA and GARCH models, as listed in Table 6.8. The test is implemented with the mean squared forecast error (MSFE) and the mean absolute forecast error (MAFE). The significance of these tests are denoted by ** (1% significance) and * (5% significance). The true daily volatility is proxied by the realized daily volatility.

Comparison	Metric	
	MSFE	MAFE
20-DAY REALIZED VOLATILITY BASED ON		
1-minute Return Frequency		
ARMA v. GARCH	-7.3262**	-6.4912**
5-minute Return Frequency		
ARMA v. GARCH	-2.9675**	-2.9604**
10-minute Return Frequency		
ARMA v. GARCH	-0.7763	-0.7813
15-minute Return Frequency		
ARMA v. GARCH	-0.3678	-0.3652
30-minute Return Frequency		
ARMA v. GARCH	-0.6750	-0.0219
20-day Return Frequency		
ARMA v. GARCH	-1.9557*	-2.0458*

**Table 6.10: Results of test for "Correct Unconditional Coverage"
Daily GARCH Models – Estimated Based on End-of-Day Returns (99% VaR)**

This table contains the results of the test for "correct unconditional coverage" in the failure series (99% VaR) estimation of the KLCI returns based on the out-of-sample VaR forecasts produced by the models below. The first column gives the names of the models, the second column gives the likelihood ratio statistic for the unconditional coverage and the third column reports the probability of success with the null hypothesis, $H_0 : \hat{f} = \alpha$ (equation 5.9) for 99% VaR coverage. The likelihood ratio statistic is given by: $LR_{uc} = -2\ln\{L(p)/L(\hat{\pi})\}$. The variables for the likelihood ratio statistic are defined in the main text. The true daily volatility is proxied by the realized daily volatility. The significance of these tests is denoted by ** (1% significance).

	Observed $\hat{f}$	Likelihood Ratio Statistic	*p*-value
<u>**Daily GARCH Models**</u>			
GARCH	0.0100	0.0000	1.0000
TGARCH	0.0100	0.0000	1.0000
EGARCH	0.0100	0.0000	1.0000

**Table 6.11: Results of test for "Independence"
Daily GARCH Models – Estimated Based on End-of-Day Returns (99% VaR)**

This table contains the results of the test for "independence" in the failure series (99% VaR) estimation of the KLCI returns based on the out-of-sample VaR forecasts produced by the models below. For each model, an OLS regression as given in equation (5.18) is performed. The first column gives the names of the models, the second column gives the estimated F-statistics of the hypothesis specified in (5.19) and the final column reports the corresponding *p*-values. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

	F-statistic	*p*-value
<u>**Daily GARCH Models**</u>		
GARCH	0.0098	0.9213
TGARCH	0.0098	0.9213
EGARCH	0.0098	0.9213

**Table 6.12: Results of test for "Correct Conditional Coverage"
Daily GARCH Models – Estimated Based on End-of-Day Returns (99% VaR)**

This table contains the results of the test for "correct conditional coverage" in the failure series (99% VaR) estimation of the KLCI returns based on the out-of-sample VaR forecasts produced by the models below. For each model, an OLS regression as given in equation (5.18) is performed. The first column gives the names of the models, the second column gives the estimated F-statistics of the hypothesis specified in (5.17) and the final column reports the corresponding *p*-values. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

	F-statistic	***p*-value**
<u>Daily GARCH Models</u>		
GARCH (1,1)	0.0051	0.9949
TGARCH (1,1)	0.0051	0.9949
EGARCH (1,1)	0.0051	0.9949

Table 6.13: Results of test for "Correct Unconditional Coverage"
Daily GARCH Models – Estimated Based on End-of-Day Returns (95% VaR)

This table contains the results of the test for "correct unconditional coverage" in the failure series (95% VaR) estimation of the KLCI returns based on the out-of-sample VaR forecasts produced by the models below. The first column gives the names of the models, the second column gives the likelihood ratio statistic for the unconditional coverage and the third column reports the probability of success with the null hypothesis, $H_0 : \hat{f} = \alpha$ (equation 5.9) for 95% VaR coverage. The likelihood ratio statistic is given by: $\text{LR}_{uc} = -2\ln\{L(p)/L(\hat{\pi})\}$. The variables for the likelihood ratio statistic are defined in the main text. The true daily volatility is proxied by the realized daily volatility. The significance of these tests is denoted by * (5% significance).

	Observed $\hat{f}$	**Likelihood Ratio Statistic**	***p*-value**
<u>Daily GARCH Models</u>			
GARCH (1,1)	0.0200	2.4286	0.1191
TGARCH (1,1)	0.0200	2.4286	0.1191
EGARCH (1,1)	0.0200	2.4286	0.1191

Table 6.14: Results of test for "Independence"
Daily GARCH Models – Estimated Based on End-of-Day Returns (95% VaR)

This table contains the results of the test for "independence" in the failure series (95% VaR) estimation of the KLCI returns based on the out-of-sample VaR forecasts produced by the models below. For each model, an OLS regression as given in equation (5.18) is performed. The first column gives the names of the models, the second column gives the estimated F-statistics of the hypothesis specified in (5.19) and the final column reports the corresponding *p*-values. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

	F-statistic	***p*-value**
<u>Daily GARCH Models</u>		
GARCH (1,1)	0.2476	0.6199
TGARCH (1,1)	0.2476	0.6199
EGARCH (1,1)	0.2476	0.6199

Table 6.15: Results of test for "Correct Conditional Coverage"
Daily GARCH Models – Estimated Based on End-of-Day Returns (95% VaR)

This table contains the results of the test for "correct conditional coverage" in the failure series (95% VaR) estimation of the KLCI returns based on the out-of-sample VaR forecasts produced by the models below. For each model, an OLS regression as given in equation (5.18) is performed. The first column gives the names of the models, the second column gives the estimated F-statistics of the hypothesis specified in (5.17) and the final column reports the corresponding *p*-values. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

	F-statistic	***p*-value**
Daily GARCH Models		
GARCH (1,1)	2.1972	0.1166
TGARCH (1,1)	2.1972	0.1166
EGARCH (1,1)	2.1972	0.1166

Table 6.16: Results of test for "Correct Unconditional Coverage"
Daily ARMA and *RiskMetrics* Models (99% VaR)

This table contains the results of the test for "correct unconditional coverage" in the failure series (99% VaR) estimation of the KLCI returns based on the out-of-sample VaR forecasts produced by the models below. The first column gives the names of the models, the second column gives the likelihood ratio statistic for the unconditional coverage and the third column reports the probability of success with the null hypothesis, $H_0 : \hat{f} = \alpha$ (equation 5.9) for 99% VaR coverage. The likelihood ratio statistic is given by: $LR_{uc} = -2\ln\{L(p)/L(\hat{\pi})\}$. The variables for the likelihood ratio statistic are defined in the main text. The true daily volatility is proxied by the realized daily volatility. The significance of these tests is denoted by ** (1% significance).

	Observed $\hat{f}$	Likelihood Ratio Statistic	*p*-value
DAILY REALIZED VOLATILITY			
1-minute Return Frequency			
ARMA	0.0300	2.6324	0.1047
5-minute Return Frequency			
ARMA	0.0200	0.7827	0.3763
10-minute Return Frequency			
ARMA	0.0200	0.7827	0.3763
15-minute Return Frequency			
ARMA	0.0200	0.7827	0.3763
30-minute Return Frequency			
ARMA	0.0100	0.0000	1.0000
One-day Return Frequency/Daily Squared Returns			
ARMA			
EGARCH	0.0100	0.0000	1.0000
RiskMetrics	0.0500	8.2582**	0.0040

Table 6.17: Results of test for "Independence"
Daily ARMA and *RiskMetrics* Models (99% VaR)

This table contains the results of the test for "independence" in the failure series (99% VaR) estimation of the KLCI based on the out-of-sample VaR forecasts produced by the models below. For each model, an OLS regression as given in equation (5.18) is performed. The first column gives the names of the models, the second column gives the estimated F-statistics of the hypothesis specified in (5.19) and the final column reports the corresponding *p*-values. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

	F-statistic	*p*-value
DAILY REALIZED VOLATILITY		
1-minute Return Frequency		
ARMA	0.0100	0.9205
5-minute Return Frequency		
ARMA	0.0099	0.9209
10-minute Return Frequency		
ARMA	0.0099	0.9209
15-minute Return Frequency		
ARMA	0.0099	0.9209
30-minute Return Frequency		
ARMA	0.0098	0.9213
One-day Return Frequency/Daily Squared Returns		
ARMA	0.0098	0.9213
EGARCH	0.0098	0.9213
RiskMetrics	0.0102	0.9196

Table 6.18: Results of test for "Correct Conditional Coverage"
Daily ARMA and *RiskMetrics* Models (99% VaR)

This table contains the results of the test for "correct conditional coverage" in the failure series (99% VaR) estimation of the KLCI returns based on the out-of-sample VaR forecasts produced by the models below. For each model, an OLS regression as given in equation (5.18) is performed. The first column gives the names of the models, the second column gives the estimated F-statistics of the hypothesis specified in (5.17) and the final column reports the corresponding *p*-values.
The significance of these tests is denoted by ** (1% significance) & * (5% significance).

	F-statistic	*p*-value
DAILY REALIZED VOLATILITY		
1-minute Return Frequency		
ARMA	0.7284	0.4853
5-minute Return Frequency		
ARMA	0.2758	0.7596
10-minute Return Frequency		
ARMA	0.2758	0.7596
15-minute Return Frequency		
ARMA	0.2758	0.7596
30-minute Return Frequency		
ARMA	0.0051	0.9949
One-day Return Frequency/Daily Squared Returns		
ARMA		
EGARCH	0.0051	0.9949
RiskMetrics	1.8017	0.1705

Table 6.19: Results of test for "Correct Unconditional Coverage"
Daily ARMA and *RiskMetrics* Models (95% VaR)

This table contains the results of the test for "correct unconditional coverage" in the failure series (95% VaR) estimation of the KLCI returns based on the out-of-sample VaR forecasts produced by the models below. The first column gives the names of the models, the second column gives the likelihood ratio statistic for the unconditional coverage and the third column reports the probability of success with the null hypothesis, $H_0 : \hat{f} = \alpha$ (equation 5.9) for 95% VaR coverage. The likelihood ratio statistic is given by: $LR_{uc} = -2\ln\{L(p)/L(\hat{\pi})\}$. The variables for the likelihood ratio statistic are defined in the main text. The true daily volatility is proxied by the realized daily volatility. The significance of these tests is denoted by * (5% significance).

	Observed $\hat{f}$	Likelihood Ratio Statistic	*p*-value
DAILY REALIZED VOLATILITY			
1-minute Return Frequency			
ARMA	0.0900	2.7510	0.0971
5-minute Return Frequency			
ARMA	0.0700	0.7530	0.3855
10-minute Return Frequency			
ARMA	0.0700	0.7530	0.3855
15-minute Return Frequency			
ARMA	0.0600	0.1984	0.6564
30-minute Return Frequency			
ARMA	0.0600	0.1984	0.6564
One-day Return Frequency/Daily Squared Returns			
ARMA	0.0200	2.4286	0.1191
EGARCH	0.0200	2.4286	0.1191
RiskMetrics	0.0800	1.6158	0.2036

Table 6.20: Results of test for "Independence"
Daily ARMA and *RiskMetrics* Models (95% VaR)

This table contains the results of the test for "independence" in the failure series (95% VaR) estimation of the KLCI returns based on the out-of-sample VaR forecasts produced by the models below. For each model, an OLS regression as given in equation (5.18) is performed. The first column gives the names of the models, the second column gives the estimated F-statistics of the hypothesis specified in (5.19) and the final column reports the corresponding *p*-values. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

	F-statistic	*p*-value
DAILY REALIZED VOLATILITY		
1-minute Return Frequency		
ARMA	3.2322	0.0753
5-minute Return Frequency		
ARMA	0.9055	0.3437
10-minute Return Frequency		
ARMA	0.9055	0.3437
15-minute Return Frequency		
ARMA	0.2592	0.6118
30-minute Return Frequency		
ARMA	0.2592	0.6118
One-day Return Frequency/Daily Squared Returns		
ARMA	0.2476	0.6199
EGARCH	0.2476	0.6199
RiskMetrics	0.5949	0.4424

**Table 6.21: Results of test for "Correct Conditional Coverage"
Daily ARMA and *RiskMetrics* Models (95% VaR)**

This table contains the results of the test for "correct conditional coverage" in the failure series (95% VaR) estimation of the KLCI returns based on the out-of-sample VaR forecasts produced by the models below. For each model, an OLS regression as given in equation (5.18) is performed. The first column gives the names of the models, the second column gives the estimated F-statistics of the hypothesis specified in (5.17) and the final column reports the corresponding p-values. The significance of these tests is denoted by ** (1% significance) and * (5% significance).

	F-statistic	*p*-value
DAILY REALIZED VOLATILITY		
1-minute Return Frequency		
ARMA	2.0365	0.1360
5-minute Return Frequency		
ARMA	0.6125	0.5441
10-minute Return Frequency		
ARMA	0.6125	0.5441
15-minute Return Frequency		
ARMA	0.2989	0.7423
30-minute Return Frequency		
ARMA	0.2989	0.7423
One-day Return Frequency/Daily Squared Returns		
ARMA	2.1972	0.1166
EGARCH	2.1972	0.1166
RiskMetrics	0.7335	0.4829

Figure 6.1: Daily Realized Volatility and Daily Squared Returns

The chart below compares the plots of the five daily realized volatilities and the daily squared returns for the 406-day sample period.
DRV 1-min = Daily Realized Volatility with 1-minute return sampling frequency
DRV 5-min = Daily Realized Volatility with 5-minute return sampling frequency
DRV 10-min = Daily Realized Volatility with 10-minute return sampling frequency
DRV 15-min = Daily Realized Volatility with 15-minute return sampling frequency
DRV 30-min = Daily Realized Volatility with 30-minute return sampling frequency

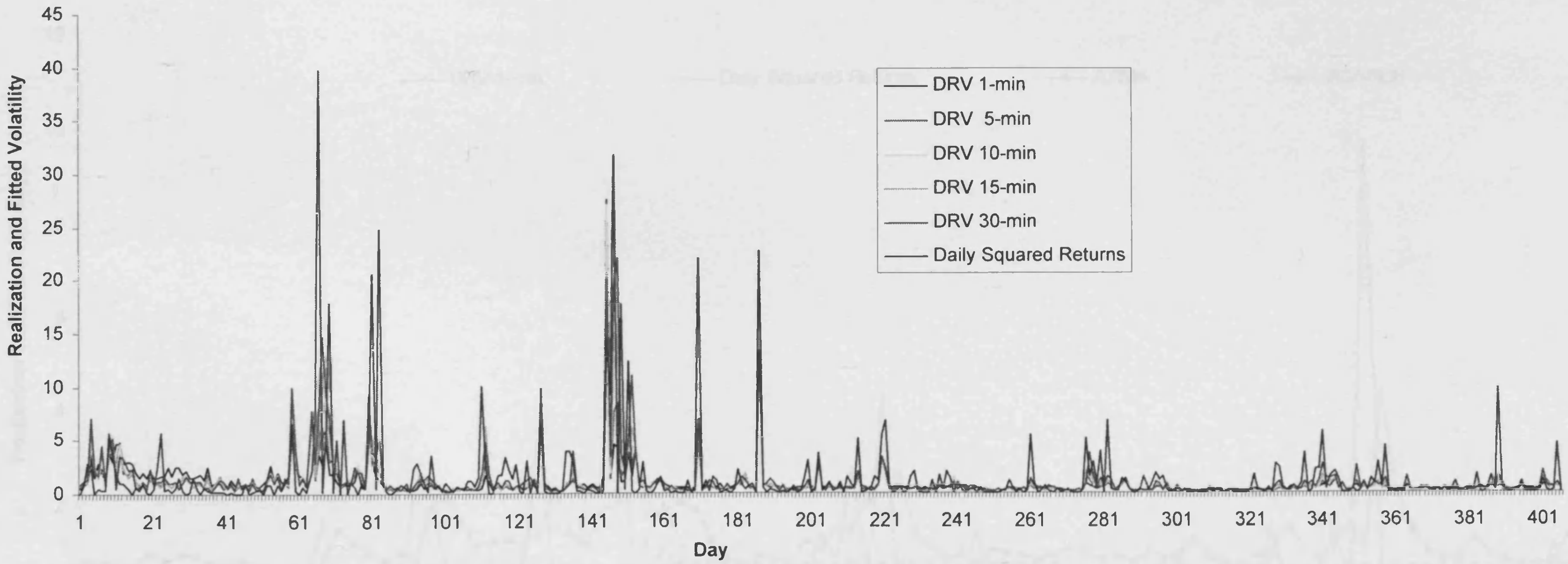


Figure 6.2: Forecasting Performance of Daily Volatility Models

The chart below compares the plots of the daily realized volatility, the daily squared returns and the forecasted return volatilities of the best performing volatility models for the 100-Day Out-of-sample forecasting period.

DRV 1-min = Daily Realized Volatility with 1-minute return sampling frequency

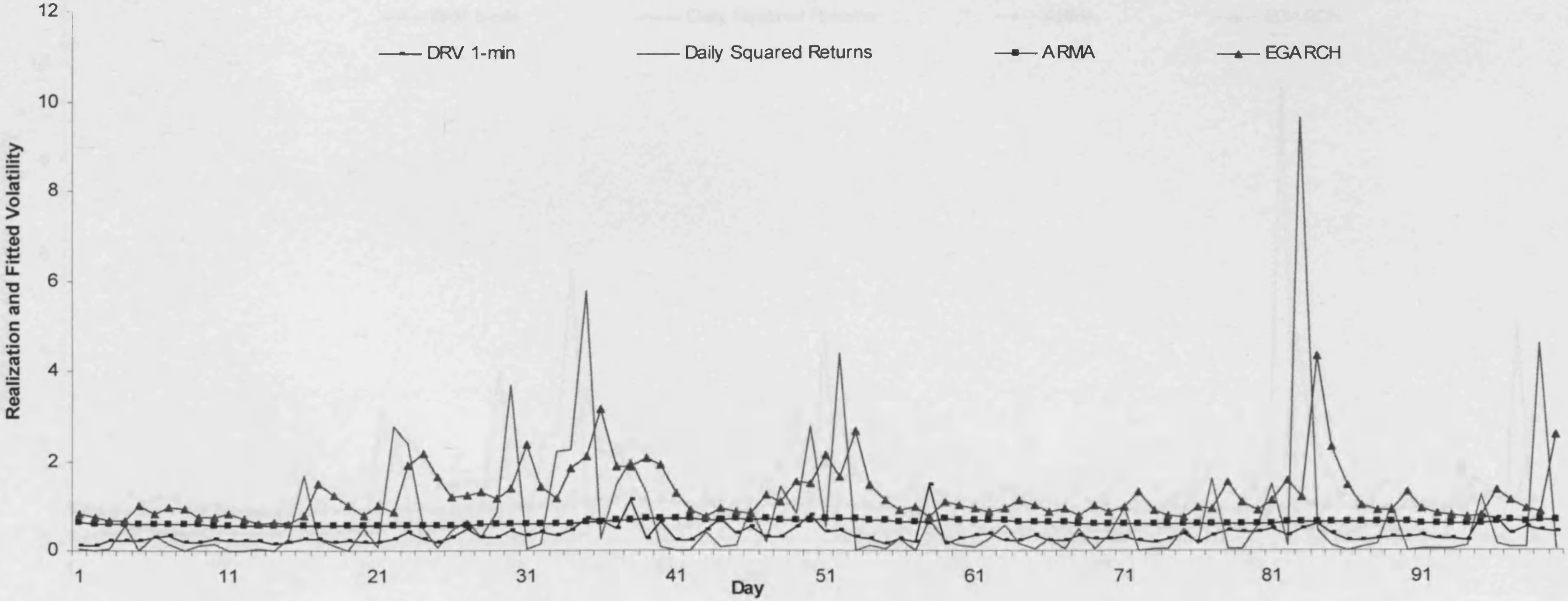


Figure 6.3: Forecasting Performance of Daily Volatility Models

The chart below compares the plots of the daily realized volatility, the daily squared returns and the forecasted return volatilities of the best performing volatility models for the 100-Day Out-of-sample forecasting period.

DRV5-min = Daily Realized Volatility with 5-minute return sampling frequency

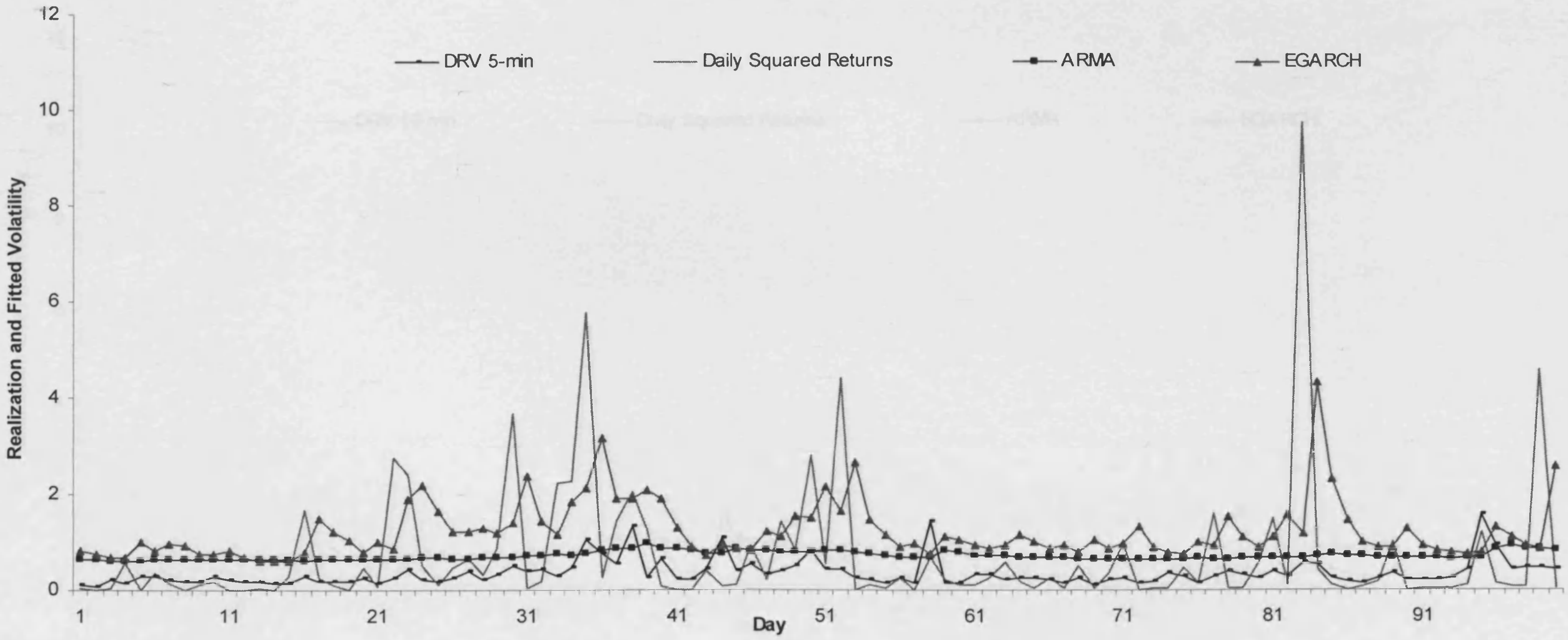


Figure 6.4: Forecasting Performance of Daily Volatility Models

The chart below compares the plots of the daily realized volatility, the daily squared returns and the forecasted return volatilities of the best performing volatility models for the 100-Day Out-of-sample forecasting period.

DRV10-min = Daily Realized Volatility with 10-minute return sampling frequency

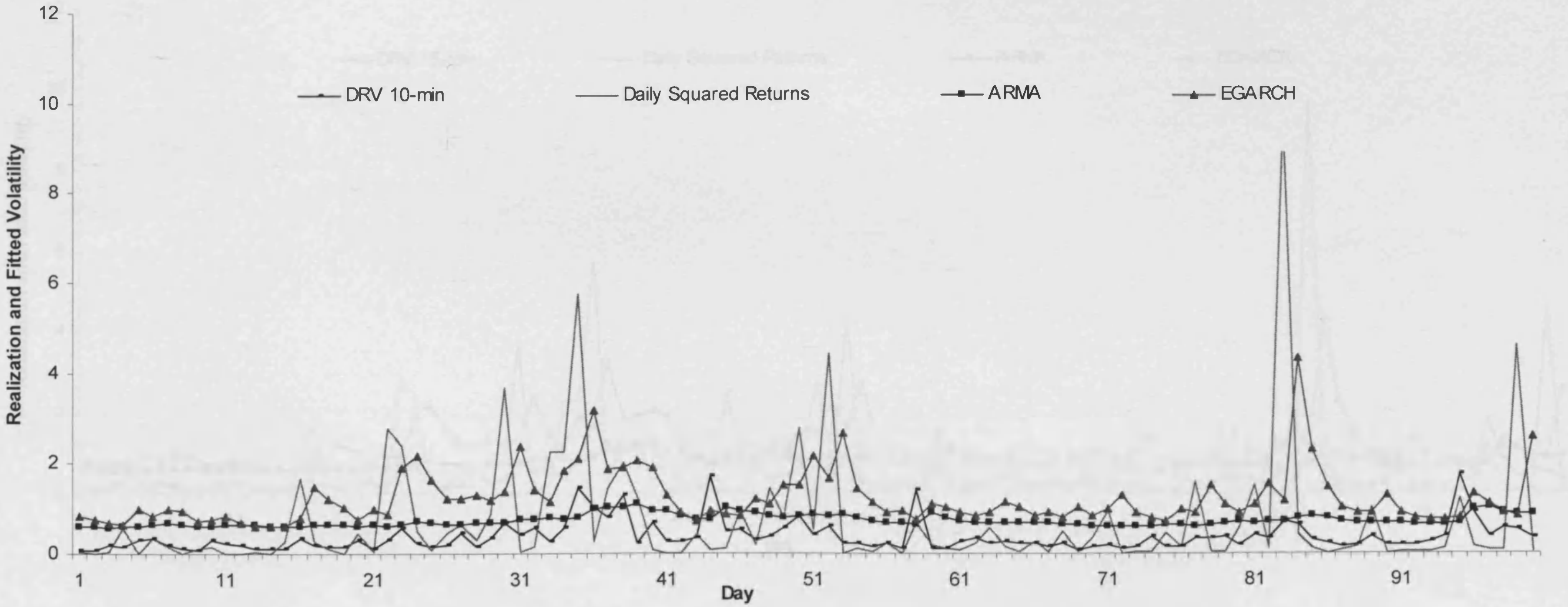


Figure 6.5: Forecasting Performance of Daily Volatility Models

The chart below compares the plots of the daily realized volatility, the daily squared returns and the forecasted return volatilities of the best performing volatility models for the 100-Day Out-of-sample forecasting period.

DRV15-min = Daily Realized Volatility with 15-minute return sampling frequency

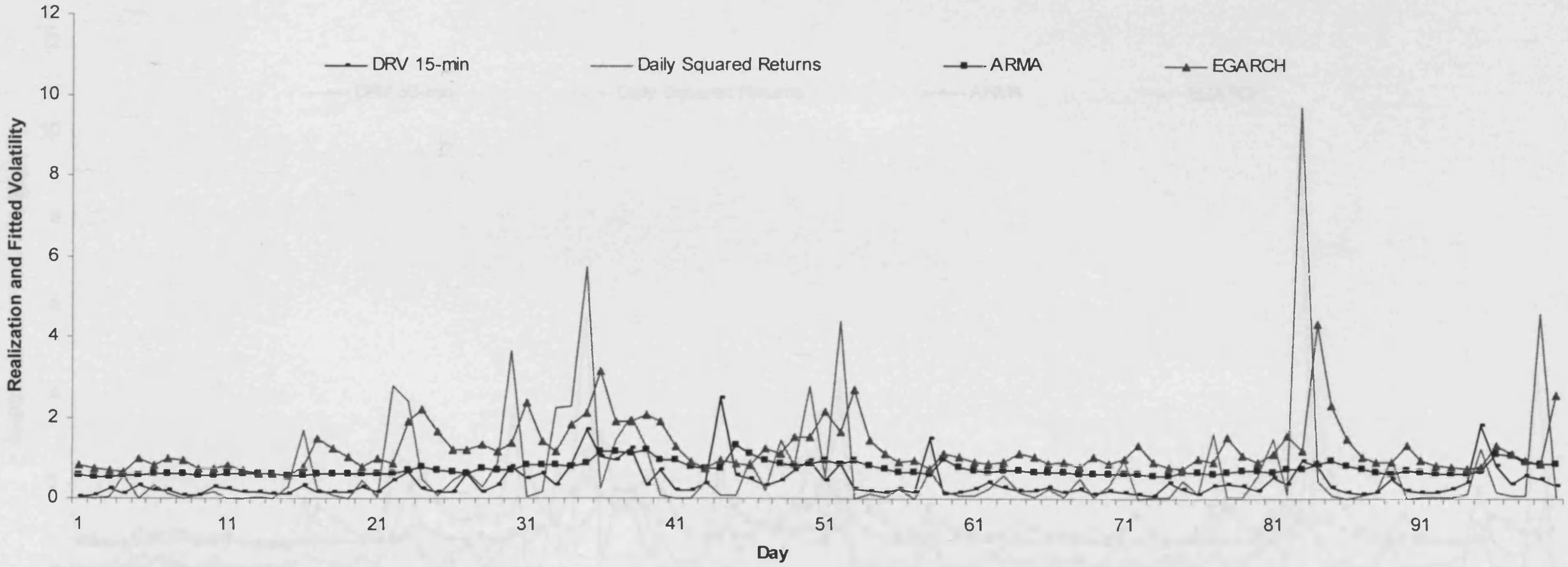


Figure 6.6: Forecasting Performance of Daily Volatility Models

The chart below compares the plots of the daily realized volatility, the daily squared returns and the forecasted return volatilities of the best performing volatility models for the 100-Day Out-of-sample forecasting period.

DRV 30-min = Daily Realized Volatility with 30-minute return sampling frequency

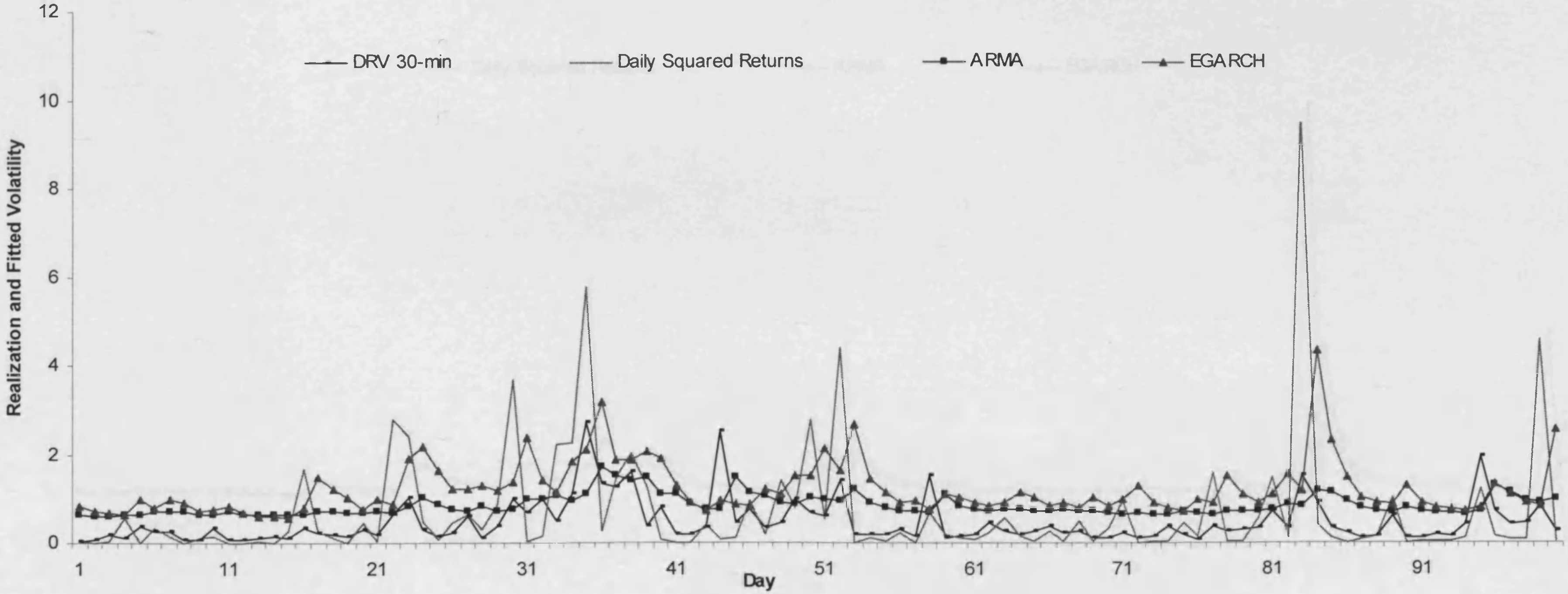


Figure 6.7: Forecasting Performance of Daily Volatility Models

The chart below compares the plots of the daily squared returns (One-day Frequency Realized Volatility) and the forecasted return volatilities of the best performing ARMA and GARCH models for the 100-Day Out-of-sample forecasting period.

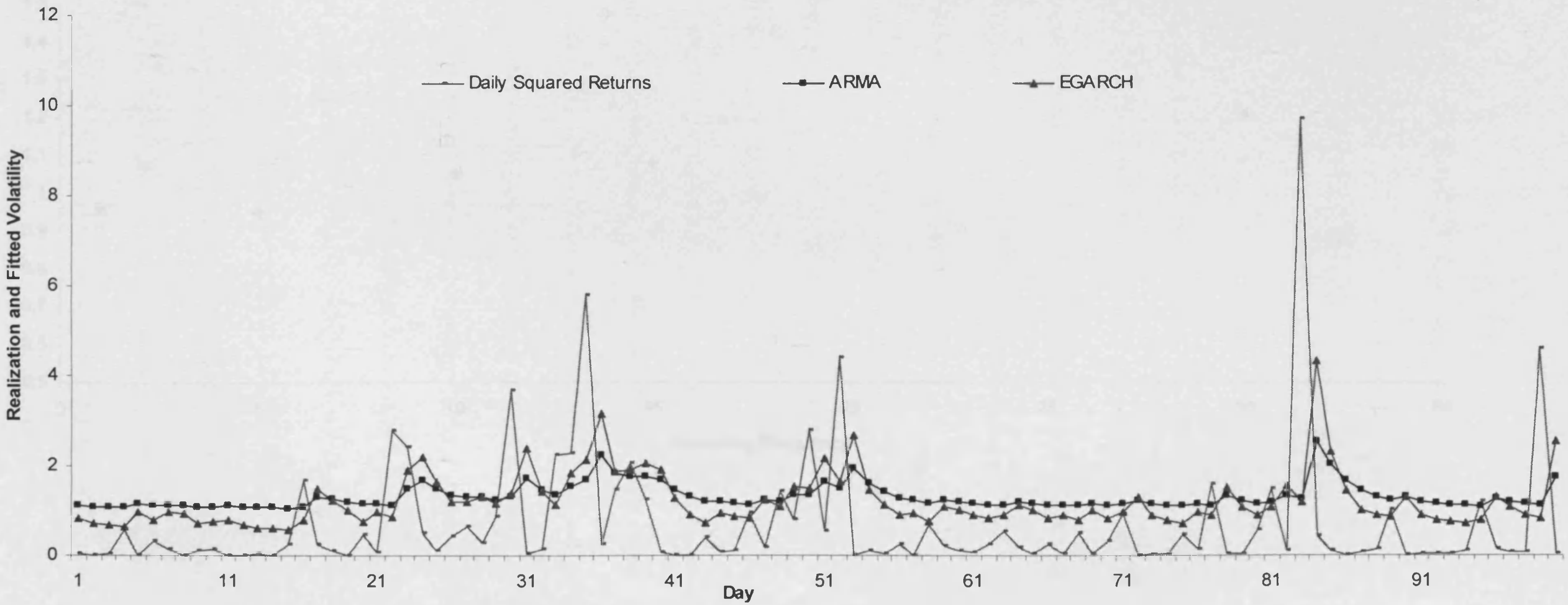
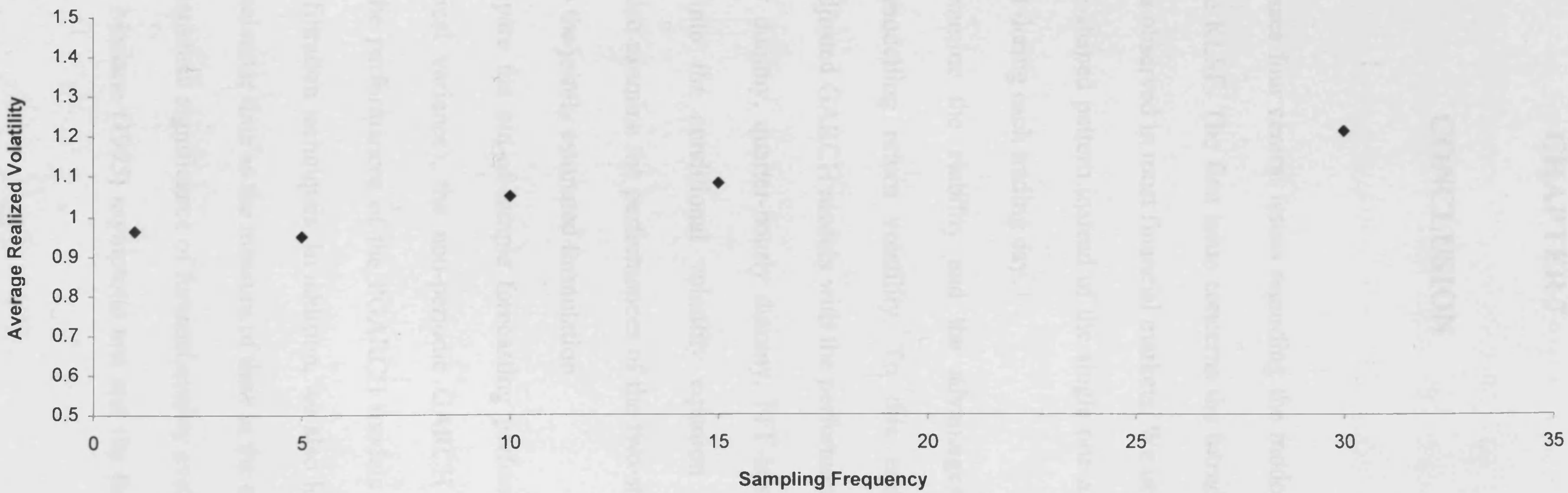


Figure 6.8: Representative Volatility Signature Plots

The chart below shows the plots of the average daily realized volatilities against the five intraday sampling frequencies used in the study: 1-minute, 5-minute, 10-minute, 15-minute and 30-minute.



CHAPTER 7

CONCLUSION

7.0 Introduction

This thesis addresses four central issues regarding the modelling and forecasting of return volatility on the KLSE. The first issue concerns the intraday U-shaped pattern for return volatility that is observed in most financial markets. We investigate whether the KLSE shows a double U-shaped pattern instead of the single one as a result of the dual trading sessions observed during each trading day.

Secondly, we examine the viability and the advantages of controlling for periodicity effects in modelling return volatility. To this end, we compare the performances of the unadjusted GARCH models with the performances of the PGARCH models with half-hourly dummy, quarter-hourly dummy, FFF-based and spline-based variables incorporated into the conditional volatility equation to account for the periodicity effects. We also examine the performances of the two-step filtration method, which is an alternative to the jointly estimated formulation.

Thirdly, we compare the out-of-sample forecasting performances of the naive model (based on historical variance), the non-periodic GARCH models (unadjusted GARCH models) with the performances of the PGARCH models using both the joint estimation and two-step filtration techniques. In addition, we also look at the impacts of using business time and calendar time as the measure of time in the estimation process. In order to emphasize the statistical significance of forecast quality evaluation, we make use of both the Diebold and Mariano (1995) asymptotic test and the forecast encompassing test of Harvey, Leybourne and Newbold (1998) to evaluate the predictive accuracy of available alternative forecasts. Following this, the economic significance of the forecast is evaluated through the performances of the various VaR measures. We assess the adequacy of VaR models constructed from these GARCH forecasts using the framework for interval forecast evaluation developed by Christoffersen (1998) and the regression-based tests of Clement and Taylor (2003).

Finally, the thesis examines the out-of-sample forecasting performance of the ARMA model, which is used to estimate the daily realized volatility measures proposed by Andersen, Bollerslev, Diebold and Labys (2001, 2003). We then evaluate the out-of-sample forecasting performances of both the ARMA and the daily GARCH-based models using the various daily realized volatility measures as the proxy for the true daily volatility. In addition, we assess the adequacy of the VaR models constructed from the ARMA and the best performing daily GARCH forecasts using the various adequacy tests described earlier.

7.1 Conclusion

We find that a reasonable double U-shaped return volatility pattern does exist for the KLCI returns for the period under study. This is shown both by the plots of the absolute returns and the standard deviation of returns during the trading day. This periodicity effect is also found to be statistically significant when we apply the Wald coefficient restriction tests. The plots of the best-fitted GARCH models from each of the thirteen volatility estimating approaches appear to model sufficiently the observed double

U-shaped pattern. The results are consistent with similar findings documented for financial markets that are closed for the lunch break period.

We find that there is a strong case for periodicity adjusting methods in modelling returns volatility. The results indicate that superior model fit is obtained for the GARCH models that are jointly estimated and the models that are estimated using the two-step filtration techniques when compared against the non-periodic (unadjusted) GARCH models. In particular, the PGARCH models that incorporate the half-hourly dummy variables in the conditional variance equation produce the best results. The results for the two-step filtration models are also encouraging, and they provide a serious alternative to the more computationally expensive method provided by the jointly estimated PGARCH models. We find that using half-hourly dummy variables gives better results compared to the alternative FFF-based variables. The same could be said for the spline-based variables. They occupy the second, the third, the fifth and the sixth positions for the best performing modelling approach. This indicates that the spline-based estimation techniques do provide a superior modelling approach, at least for the Malaysian market.

The superior results of the PGARCH model fits are also translated into superior out-of-sample forecasting performances. Specifically, PGARCH forecasts give smaller forecast error statistics than the standard unadjusted GARCH equivalents. We find that the quality of forecasts of the PGARCH-based models is superior to the quality of the naive and the standard GARCH forecasts (both the results of the Diebold and Mariano (1995) asymptotic test and the encompassing test of Harvey, Leybourne and Newbold (1998), using the MSFE and MAFE metrics, are statistically significant). Therefore, it is clearly desirable to adjust and control for periodicity effects, in terms of producing superior forecast quality, and this does add additional informational content to the quality of forecasts. However, we find that VaR models constructed based on high-frequency forecasted data are sensitive to the effects of the fat tails of the series distribution. Assuming normal distribution, the VaR produced by the available GARCH forecasts fails miserably when the coverage tests are applied at both the 99% and the 95% levels.

Finally, we demonstrate the superiority of the daily realized volatility measure, based on the intraday summation of squared returns, over the daily squared returns measure. We observe that the daily GARCH models produce superior forecasting performances when the daily realized volatility measures are used as the proxy for the true daily volatility. This is consistent with the results of previous studies. The results, therefore, suggest that in order to optimise the application of the GARCH models, one should consider using the integrated realized volatility measure as the proxy for the actual volatility. Based on the volatility signature plot, we recommend that the most appropriate sampling frequency for the daily realized volatility among the five different return sampling frequencies examined in this study is the 5-minute return frequency. This is certainly consistent with the 5-minute sampling frequency often used and cited for the developed capital markets.

In addition, we also find that the ARMA model used to model the daily realized volatilities produce superior forecasting performances compared to the daily GARCH models when the same daily realized volatility is used as the benchmark. The results, however, are rather mixed when the daily squared returns are considered. Here, for the MAFE statistics, the daily GARCH models are superior to the ARMA model, while the opposite is true when the MSFE statistics are taken into account. The results of the

Diebold and Mariano (1995) asymptotic tests are also significant and suggest that in all paired comparisons of alternative forecasts, the quality of the ARMA forecasts is superior to the quality of the GARCH forecasts, with the exception of forecast evaluation using the one-day frequency realized volatility, where we find that the quality of both the ARMA and the GARCH forecasts is the same when the MSFE metric is applied. Finally, there are some positive results for VaR models based on the ARMA and the daily GARCH forecasts. The GARCH-based VaR models appear to have the required coverage properties. In addition, the results indicate that the ARMA model that were sampled at the 1-minute, 5-minute, 10-minute, 15-minute, 30-minute return and one-day sampling frequencies easily satisfy the required VaR criteria. The same could be said for the *RiskMetrics* model, which passes all the correct conditional coverage tests conclusively. Turning to the most appropriate sampling frequency for the daily integrated realized volatility, we find that the 5-minute return sampling frequency provides the lowest average daily volatility and therefore produces the best proxy for the true daily volatility.

REFERENCES

Admati, A. & P. Pfleiderer. (1988). A theory of intraday patterns: Volume and price variability. Review of Financial Studies 1, 3-40.

Aggarwal, R. & E. Gruca. (1993). Intra-day trading patterns in the equity options markets. Journal of Financial Research 16, 285-297.

Akgiray, V. (1989). Conditional heteroskedasticity in time series of stock returns: evidence and forecasts. Journal of Business 62, 55-80.

Andersen, T. & T. Bollerslev. (1997). Intraday periodicity and volatility persistence in financial markets. Journal of Empirical Finance 4, 115-118.

Andersen, T. & T. Bollerslev. (1998a). DM-Dollar volatility: Intraday activity patterns, macroeconomic announcements and longer run dependencies. Journal of Finance 53, 219-265.

Andersen, T. & T. Bollerslev. (1998b). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. International Economic Review 39, 885-906.

Andersen, T. G. Bollerslev. T. & J. Cai. (2000). Intraday and interday volatility in the Japanese stock market. Journal of International Financial Markets, Institutions and Money 10, 107-130.

Andersen, T.G., Bollerslev, T., Christoffersen, P. F. & F. X. Diebold, (2005). Volatility forecasting. NBER Working Paper Series 11188.

Andersen, T.G., Bollerslev, T., Diebold, F. X. & H. Ebens. (2001). The distribution of stock return volatility. Journal of Financial Economics 61, 43-76.

Andersen, T.G., Bollerslev, T., Diebold, F. X. & P. Labys. (2001). The distribution of realized exchange rate volatility. Journal of the American Statistical Association 96, 42-55.

Andersen, T.G., Bollerslev, T., Diebold, F. X. & P. Labys. (2003). Modeling and forecasting realized volatility. Econometrica 71(2), 579-626.

Areal, N. M. P. C. & S. J. Taylor. (2002). The realised volatility of FTSE-100 future prices. Journal of Futures Markets 22, 627-648.

Ariff, M., Shamsher, M. & M. N. Annuar. (1998). Stock pricing in Malaysia. Universiti Putra Malaysia Press.

Bachelier, L. (1900). Trans. James Boness. Theory of speculation. In Cootner (1964), 17-78.

Baillie, R. & T. Bollerslev. (1991). Intra-day and inter-market volatility in foreign exchange rates. Review of Economic Studies 58, 565-585.

Bali, T. G. & Y. Lu. (2005). Forecasting stock market volatility: The empirical performance of SGED-GARCH, implied, and realized volatility models. Seminar Paper. The Graduate Center of the City University of New York.

Barndorff-Nielsen, O. L. & N. Shephard. (2002a). Econometric analysis of realised volatility and its use in estimating Stochastic Volatility models. Journal of the Royal Statistical Society 64, 253-280.

Barndorff-Nielsen, O. L. & N. Shephard. (2002b). Estimating quadratic variation using realised variance. Journal of Applied Econometrics 17, 457-477.

Basle Committee. (1996). Amendment to the Capital Accord to Incorporate Market Risks.

Basle Committee. (1998). Amendment to the Capital Accord to Incorporate Core Principles for Effective Banking Supervision.

Bera, A. & M. Higgins. (1993). A survey of ARCH models: Properties, estimation and testing. Journal of Economic Surveys 7, 305-366.

Bera, A. K. & M. L. Higgins. (1997). ARCH and bilinearity as competing models for nonlinear dependence. Journal of Business & Economic Statistics, 15, 43–50.

Berkowitz, J. & J. O'Brien. (2002). How accurate are Value-at-Risk models at commercial banks? Journal of Finance 57 (3), 1093-1111.

Bildik, R. (2000). Intra-day seasonalities on stock returns: Evidence from the Turkish stock market. Paper presented at the Annual Meeting of the European Financial Management Association, Athens.

Black, F. (1976). The pricing of commodity contracts. Journal of Financial Economics 3, 167-179.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. Journal of Econometrics 31, 307-327.

Bollerslev, T., Chou, R. Y. & K. P. Kroner. (1992). ARCH modeling in finance: a review of theory and empirical evidence. Journal of Econometrics 52, 5-59.

Bollerslev, T. & I. Domowitz. (1993). Trading patterns and prices in the interbank foreign exchange market. Journal of Finance 48, 1421-1443.

Bollerslev, T., Engle, R. & D. Nelson. (1994). ARCH models, in R. F. Engle and D. McFadden (eds). Handbook of Econometrics, 4 (North-Holland, Amsterdam).

Bollerslev, T. & E. Ghysels. (1996). Periodic autoregressive conditional heteroscedasticity. Journal of Business and Economic Statistics 14, 139-151.

Brailsford, T. J. & R. W. Faff. (1996). An evaluation of volatility forecasting techniques. Journal of Banking and Finance 20, 419-438.

Bredin, D. & S. Hyde. (2004). FOREX risk: Measurement and evaluation using Value-at-Risk. Journal of Business Finance & Accounting 31 (9-10), 1389-1417.

Brock, W. & A. Kleidon. (1992). Periodic market closure and trading volume. Journal of Economic Dynamics and Control 16, 451-489.

Brooks, C. (2002). Introductory econometrics for finance. Cambridge University Press.

Brooks, C. & G. Persand. (2003). Volatility forecasting for risk management. Journal of Forecasting 22, 1-22.

Campbell, J. Y., Lo, A. W. & A. C. Mackinlay. (1997). The econometrics of financial markets. Princeton University Press.

Cao, C. Q. & R. S. Tsay. (1992). Nonlinear time-series analysis of stock volatilities. Journal of Applied Econometrics Supplement 1, 165-185.

Chan, K., Chan, K. C. & G. A. Karolyi. (1991). Intraday volatility in the stock index and stock index futures market. Review of Financial Studies 4, 657-684.

Chang, R., Fukuda, T., Rhee, S. & M. Takano. (1993). Intraday and intraweek behaviour of the TOPIX. The Pacific-Basin Finance Journal 1, 67-95.

Cheung, Y., Ho, R., Pope, P. & P. Draper. (1994). Intraday stock return volatility: The Hong Kong evidence. Pacific-Basin Finance Journal 2, 261-276.

Choe, H. & H. S. Shin. (1993). An analysis of interday and intraday return volatility: Evidence from the Korea Stock Exchange. The Pacific-Basin Finance Journal 1, 175-188.

Chong, C. W., Ahmad, M. I. & M. Y. Abdullah. (1999). Performance of GARCH models in forecasting stock market volatility. Journal of Forecasting 18, 333-343.

Chou, R. Y. (1988). Volatility persistence and stock valuation: Some empirical evidence using GARCH. Journal of Applied Econometrics 3, 279-294.

Christoffersen, P. F. (1998). Evaluating interval forecasts. International Economic Review 39, 841-862.

Christoffersen, P. F., Diebold, F. X. & T. Schuermann. (1998). Horizon problems and extreme events in financial risk management. Economic Policy Review, Federal Reserve Bank of New York, October, 109-118.

Clements, M. & N. Taylor. (2003). Evaluating interval forecasts of high-frequency financial data. Journal of Applied Econometrics 18(4), 445-456.

Cornett, M. M., Schwarz, T. V. & A. C. Szakmary. (1995). Seasonalities and intraday return patterns in the foreign currency futures market. Journal of Banking and Finance 19, 843-869.

Cowles, A. (1933). Can stock market forecasters forecast? Econometrica 1, 309-324.

Cowles, A. (1944). Stock market forecasting. Econometrica 12, 206-214.

Cumby, R., Figlewski, S. & J. Hasbrouck. (1993). Forecasting volatilities and correlations with EGARCH models. Journal of Derivatives 1, 51-63.

Dacorogna, M. M., Muller, U. A., Nagler, R. J., Olsen, R. B. & O. V. Pictet. (1993). A geographical model for the daily and weekly seasonal volatility in the foreign exchange market. Journal of International Money and Finance 12, 413-438.

De Brouwer, G. (2002). Financial markets, institutions and integration in East Asia. Paper presented at the Asian Economic Panel, Tokyo.

Diebold, F. X. & R. S. Mariano. (1995). Comparative predictive accuracy. Journal of Business and Economic Statistics 13, 253-263.

Dimson, E. & M. Mussavian. (1998). A brief history of market efficiency. European Financial Management 4(1), 91-193.

Drost, F. C. & T. E. Nijman. (1993). Temporal aggregation of GARCH processes. Econometrica 61(4), 900-927.

Ebens, H. (1999). Realized stock index volatility. Working Paper no. 420, Department of Economics, John Hopkins University.

Engle, R.F. (1982). Autoregressive conditional heteroscedasticity, with estimates of the variance of United Kingdom inflation. Econometrica 50, 987-1007.

Engle, R.F. (2000). New frontiers for ARCH models. Journal of Applied Econometrics 17, 425-446.

Engle, R. (2001). GARCH 101: The use of ARCH/GARCH models in applied econometrics. Journal of Economic Perspectives 5, 157-168.

Engle, R. F., Lilien, D. M. & R. P. Robins. (1987). Estimating time varying risk premia in the term structure: The ARCH-M model. Econometrica 55 (3), 391-407.

Engle, R. F. & S. Manganelli. (2004). CAViaR: Conditional autoregressive Value-at-Risk by regression quantiles. Journal of Business and Economic Statistics 22 (4), 367-381.

Engle, R. F. & V. K. Ng. (1993). Measuring and testing the impact of news on volatility. Journal of Finance 48, 1749-1778.

Fama, E. F. (1963). Mandelbrot and the stable Paretian distribution. Journal of Business 36, 420-429.

Fama, E. F. (1965). The behaviour of stock market prices. Journal of Business 38, 34-105.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. Journal of Finance 25, 383-417.

Fama, E. F. & K. French. (1988). Permanent and temporary components of stock prices. Journal of Political Economy 96, 246-273.

Figlewski, S. (1997). Forecasting volatility. Financial Markets, Institutions, and Instruments. New York University Salomon Center 6(1), 1-88.

Foster, F. & S. Viswanathan. (1990). A theory of the interday variations in volume, variance, and trading costs in securities markets. Review of Financial Studies 3, 593-624.

Foster, F. & S. Viswanathan. (1993). Variations in trading volume, return volatility and trading costs: Evidence on recent price formation models. Journal of Finance 48, 187-211.

Franses, P. H. & D. van Dijk. (1996). Forecasting stock market volatility using (nonlinear) GARCH models. Journal of Forecasting 15, 229-235.

French, K.R., Schwert, G.W. & R. F. Stambaugh. (1987). Expected stock returns and volatility. Journal of Financial Economics 19, 3-30.

Gallant, A. R. (1981). On the bias in flexible functional forms and an essentially unbiased form: the Fourier flexible form. Journal of Econometrics 15, 211-245.

Gallant, A. R. (1982). Unbiased determination of production technologies. Journal of Econometrics 20, 285-323.

Gerety, M.S. & J. H. Mulherin. (1992). Trading halts and market activity: An analysis of volume at the open and the close. Journal of Finance 47(5), 1765-1784.

Giot, P. & S. Laurent. (2004). Modelling daily Value-at-Risk using realized volatility and ARCH type models. Journal of Empirical Finance 11 (3), 379-398.

Gigli, A. (2002). Quantitative finance page.[www]<URL http://www.giglia.it/MarketvolIntro.PDF/>[Accessed 9 March 2004]

Gokcan, S. (2000). Forecasting volatility of emerging stock markets: Linear versus non-linear GARCH models. Journal of Forecasting 19(6), 499-504.

Hamilton, J. D. (1994). Time series analysis. Princeton University Press, New Jersey.

Harris, L. (1986). A transaction data study of weekly and intradaily patterns in stock returns. Journal of Financial Economics 16, 99-117.

Harvey, D. I., Leybourne S. J. & P. Newbold. (1998). Tests for forecast encompassing. Journal of Business & Economic Statistics 16, 254-259.

Jain, P. & G. Joh. (1988). The dependence between hourly prices and trading volume. Journal of Financial and Qualitative Analysis 23, 269-283.

Jordan, J. V. Seale, W. E., Dinehart, S. & D. Kenyon. (1998). The intraday variability of soybean futures prices: Information and trading effects. Review of Futures Markets 7, 97-108.

Jorion, P. (1995). Predicting volatility in the foreign exchange market. Journal Finance 50(2), 507-528.

Jorion, P. (1996). Value at Risk: The new benchmark for controlling market risk. Irwin, Chicago.

J.P. Morgan/Reuters. (1996). RiskMetrics – technical document. New York 4th edn.

Keim, D. & R. Stambaugh. (1986). Predicting returns in stock and bond markets. Journal of Financial Economics 17, 357-390.

Kendall, M. (1953). The analysis of economic time series. Journal of the Royal Statistical Society, Series A 96, 11-25.

Kim, O. & R. E. Verrecchia. (1994). Liquidity and volume around earnings announcements. Journal of Accounting and Economics 17, 41-67.

Lam, H. & W. Tong. (1999). Interdaily volatility in a continuous order-driven market . Journal of Business Finance and Accounting 27, 1013-1036.

Lazarov, Z. (2004). Modeling and forecasting DAX index volatility. Bonn Econ Discussion Papers 5.

Lo, A. W. & A. C. MacKinlay. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. The Review of Financial Studies 1, 41-66.

Locke, P. R. & C. L. Sayers. (1993). Intra-day futures price volatility: Information effects and variance persistence. Journal of Applied Econometrics 8, 15-30.

Lopez, J. (1999). Methods for evaluating Value-at-Risk estimates. Federal Reserve Bank of San Francisco Economic Review 2, 3-17.

Malkiel, B. (1979). The capital formation problem in the United States. Journal of Finance 34, 291-306.

Mandelbrot, B. (1963). The variation of certain speculative prices. Journal of Business 36(4), 394-419.

Manganelli, S. & R. F. Engle. (2001). Value at risk models in finance. Working Paper 75, European Central Bank.

Martens, M. (2001). Forecasting daily exchange rate volatility using intraday returns. Journal of International Money and Finance 20, 1-23.

Martens, M., Chang, Y. C. & S. Taylor. (2002). A comparison of seasonal adjustment methods when forecasting intraday volatility. Journal of Financial Research 25, 413-423.

McMillan, D. G. & A. E. H. Speight. (2004a). Daily volatility forecasts: Reassessing the performance of GARCH models. Journal of Forecasting 23, 449-460.

McMillan, D. G. & A. E. H. Speight. (2004b). Intra-day periodicity, temporal aggregation and time-to-maturity in FTSE-100 index futures volatility. Journal of Applied Financial Economics 14, 253-263.

McMillan, D. Speight, A. & O. Gwilym. (2000). Forecasting UK stock market volatility. Journal of Applied Financial Economics 10, 435-448.

Melvin, M. & B. P. Melvin. (2003). The global transmission of volatility in the foreign exchange market. The Review of Economics and Statistics 85(3), 670-679.

Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. Journal of Financial Economics 8, 323-361.

Mohammed, Z., Fauzias, M. & Y. Othman. (1995). Intraday returns patterns of Malaysian common stock. Jurnal Pengurusan 14, 43-58.

Nelson, D. (1991). Conditional heteroskedasticity in asset returns: A new approach. Econometrica 59, 347-370.

Nelson, D. & C. Cao. (1992). Inequality constraints in the univariate GARCH model. Journal of Business & Economic Statistics 10, 229-235.

Oomen, R. (2001). Using high frequency stock market index data to calculate, model and forecast realized volatility. Manuscript, European University Institute, Department of Economics.

Pagan, A. & G. Schwert. (1990). Alternative models for conditional stock volatility. Journal of Econometrics 45, 267-290.

Pandey, A. (2003). Modeling and forecasting volatility in Indian capital markets. Indian Institute of Management Ahmedabad Working Papers.

Peterson, D. R. (1990). A transaction data study of day-of-the-week and intraday patterns in option returns. Journal of Financial Research 13, 117-131.

Pindyck, R. S. (1984). Risk, inflation and the stock market. American Economic Review 74, 335-351.

Pindyck, R. S. (1986). Risk aversion and the determinants of stock market behaviour. Working Paper 1921, National Bureau of Economic Research New York.

Pong, S., Shackleton, M.B., Taylor, S. J. & X. Xu. (2002). Forecasting Sterling/Dollar volatility; implied volatility versus long memory intraday models. Working Paper. Lancaster University.

Poon, S. & C. Granger. (2003). Forecasting volatility in financial markets: A Review. Journal of Economic Literature 41 (2), 478-539.

Poon, S. & C. Granger. (2005). Practical issues in forecasting volatility. Financial Analysts Journal 61(1), 45-56.

Porteba, J. & L. Summers. (1986). The persistence of volatility and stock market fluctuations. American Economic Review 76, 1142-1151.

Rabemananjara, R. & J. Zakoian. (1993). Threshold ARCH models and asymmetries in volatility. Journal of Applied Econometrics 8, 31-49.

Roberts, H. (1967). Statistical versus clinical prediction of the stock market. Unpublished manuscript, CRSP, University of Chicago.

Samuelson, P. (1965). Proof that properly anticipated prices fluctuate randomly. Industrial Management Review 6, 41-49.

Schwert, G. W. (1989). Why does stock market volatility change over time? Journal of Finance 44, 1115-1153.

Schwert, G. (1990). Stock market volatility. Financial Analysts Journal 46, 23-34.

Taylor, N. (2004). Modeling discontinuous periodic conditional volatility: Evidence from the commodity futures market. Journal of Futures Market 24(9), 805-834.

Taylor, S. (1986). Modelling Financial Time Series. John Wiley, New York.

Tse, Y. K. (1991). Stock returns volatility in the Tokyo Stock Exchange. Japan and the World Economy 3, 285-298.

Tse, S. H. & K. S. Tung. (1992). Forecasting volatility in the Singapore Stock Market. Asia Pacific Journal of Management 9, 1-13.

West, K. D. & D. Cho. (1995). The predictive ability of several models of exchange rate volatility. Journal of Econometrics 69(2), 367-391.

Wood, R. & McInish, T. & J. Ord. (1985). An investigation of intraday data for NYSE stocks. Journal of Finance 40, 723-741.

Yadav, P. K. & P. F. Pope. (1992). Intraweek and intraday seasonalities in stock market risk premia. Journal of Banking and Finance 16(1), 233-270.

Zakoian, J. M. (1994). Threshold heteroskedastic models. Journal of Economic Dynamics and Control 18, 931-955.