Three Essays in the Financial Economics of Conditional Volatility

Jingyi Liu
Principal Supervisor: Professor Andy Snell
2nd Supervisor: Dr. Angelica Gonzalez

Submitted for the degree of
Doctor of Philosophy
Autumn 2008

The University of Edinburgh
School of Business and Economics

The dissertation of Jingyi Liu is approved:

Chair Date

Date

Date

The University of Edinburgh

Autumn 2008

Three Essays in the Financial Economics of Conditional Volatility

Declaration

I certify that this thesis does not incorporate any material previously submitted for a degree or diploma in any University; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person where due reference is not made in the text. I declare that, except where otherwise stated, the work contained herein is my own original contribution.

Chapter 2, *Forecasting Foreign Exchange Volatility* was presented at the 26th International Symposium on Forecasting (ISF) in Santander, Spain, and the 3th Annual Meeting of Multinational Finance Society (MFS) in Edinburgh etc. It has also been accepted by more than five international conferences including the 17th Annual Asian Finance Association (Asian FA/FMA) in Auckland, New Zealand.

Chapter 3, *Can A Lucas Model with Habit Generate Realistic Conditional Volatility in Exchange Rate Returns?*, has been accepted by the 17th Annual Meeting European Financial Management Association (EFMA) in Athens, Greece.

An abridged version of Chapter 3 has been written up as a paper joint with Angelica Gonzalez entitled "*A Habit-Based Explanation to Conditional Volatility in FOREX Returns*" and submitted to the Journal of Empirical Finance (JEF) to be considered for possible publication. It is noted that some of the research in the paper was joint work with Angelic Gonzalez but throughout I remained the lead researcher and did the majority of the work.

Chapter 4, *Forecasting Volatility: Optimal Forecast Error Criterion for Utility-Based Loss Functions*, has been submitted to the 12th Conference of the Swiss Society for Financial Market Research in Geneva, Switzerland, and the 2008 Financial Research Association Meeting in Las Vegas, Nevada.

Abstract

Volatility is associated with risk and uncertainty. Financial market volatility plays an important role in investment, option pricing, risk management, and monetary policy making. Conditional volatility is one of the most prominent properties of volatility in financial markets. Nearly all empirical work in finance published this decade is involved with conditional volatility in returns.

This thesis concentrates on investigating three important questions on conditional volatility in financial markets: if volatility is forecastable, which method will provide the best forecasts? What economic behaviour is the reason behind conditional volatility, if any? What optimal statistical evaluation criterion of conditional forecasts does the economic utility maximization correspond to? The new ideas, viewpoints, methodologies and theoretical underpinning employed in this thesis endow the study on conditional volatility in financial markets with a deep and comprehensive understanding in the need for better controlling and modeling asymmetric and clustering volatility. To consider these questions, there are three main chapters composing the thesis, which can be read independently. It is emphasized that Chapter 3 plays an important role in the main strengths of this thesis.

In Chapter 2, we investigate the out-of-sample predictive ability of 73 competing time series models for the volatility of foreign exchange changes. Using the evaluation criteria of forecast accuracy and efficiency tests, we compare the out-of-sample forecasting performance of the monthly volatility of the US Dollar versus UK Sterling exchange rate from the post-Bretton Woods era to the present day. The empirical results support the stylized facts of volatility. Historical volatility models are superior to ARCH class models. However, ARCH class models take predominance where over-predictions are more heavily penalized. The various model ranks are shown to be sensitive to the error statistics used to assess the accuracy of the forecasts. The frequency of the data, the type of information used, the forecast horizon, the forecast model, and the evaluation criterion are all important variables in volatility forecasting. There is no single forecasting model suitable for all purposes.

In Chapter 3, we present a theoretical underpinning to the well established empirical stylized fact that asset returns in general and the spot foreign exchange returns in particular display predictable volatility characteristics. Adopting Moore and Roche's habit persistence version of the Lucas model we find that both the innovation in the spot foreign exchange returns and the foreign exchange returns itself follow "ARCH" style processes. Using the impulse response functions we show that the baseline simulated foreign exchange series has "ARCH" properties in the quarterly frequency that match well the "ARCH" properties of the empirical monthly estimations in Chapter 2 in that when we scale the x-axis to synchronize the monthly and quarterly responses we find similar impulse responses to one unit shock in variance. The impulse response functions for the ARCH processes we estimate "look the same" with an approximately monotonic decreasing fashion. The Lucas two-country monetary model with habit can generate realistic conditional volatility in spot foreign exchange returns.

In Chapter 4, we propose an optimal forecast error criterion for utility maximization under an option trading rule. Analysing the quadratic and exponential utility functions, which give the "utility" or "loss" of the cumulated profits from the repeated daily S&P 500 index option trade, we find that both utility cases are asymmetric and peak when

the forecast conditional variance equals the actual conditional variance (forecast error is zero). In the sense that the expected utility is a declining function of forecast error, we regress expected utility on forecast error and find that the coefficients in the regression depend on the parameters in the economic problem an investor faces, including the risk aversion parameter and the level of conditional variance. Taking the averaged form of the regression gives the approximate optimal forecast error criterion in terms solely of recognizable statistical loss functions like MAE, MSE etc. We repeat this procedure for different levels of risk aversion and study how the regression coefficients change when the risk aversion parameter changes. The empirical results show that for a more highly risk averse investor the optimal forecast error criterion is a weighted average of MAE and MSE but which weights MSE less heavily. The optimality forecast error criterion based on functions of forecast errors for utility maximization under asymmetric loss provides a simple rule for making economic and financial decisions under uncertainty.

Acknowledgements

Looking back, I'm overwhelmed by how much of a debt of gratitude I owe to those special people who helped in many ways.

My principal supervisor Professor Andy Snell provided many excellent comments and suggestions that are to the point and thought provoking and that have helped to improve the thesis substantially. His care-to-detail and conscientious scientific attitude is exemplary and infectious; and the reading of his comments is itself a learning and elevating process. I sincerely appreciate his teaching and encouragement, which inspired me to continue with my doctoral studies. His keen insight into people and cosmopolitan attitude proved invaluable. I also thank him for his constant support along the way.

Special thanks are owed to my second supervisor Dr. Angelica Gonzalez. She was generous in providing many helpful comments and continued patience. I thank her for pointing out those many errors that otherwise would have been ignored and for sharing her thoughts on many issues and for being a good "tutor" in many respects.

To my parents and sisters, who always have a special place in my mind and whose constant love, understanding and unwavering support have been a never-ending source of inspiration, I owe an enormous debt of gratitude.

To my best friends, with whom many moments of laughter and sorrows were shared and with whom many nights of chatting and discussion were passed, I'm very thankful to have all of you.

To all those around me, without the understanding of whom, the writing of this thesis – mainly away from everyday circles – would not be possible, a blanket thanks here will have to suffice.

I gratefully acknowledge the School of Business and Economics for the facilities and the staff for all its help along the way.

To my parents

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The recent rise in financial market volatility starting from US subprime mortgage crisis has increased uncertainty in the world economy. The latter part of 2007 witnessed considerable turmoil in international financial markets, especially in the exchange rate and derivative markets[1]. Volatility in most markets fluctuates at the turn of the millennium. Shifts in volatility affect investors' willingness to hold risky assets and their prices. The level of volatility in financial markets can also influence corporations' investment decisions and banks' willingness and ability to extend credit. Sharp changes in the level of financial market volatility can also be of concern to policy makers. A sudden increase in volatility might discourage major market participants from providing two-way price quotations, which in turn can reduce liquidity and trigger adverse price reactions, with potential consequences for the real economy.

The importance of financial volatility is demonstrated by the large literature it has given rise to. During the last two decades, volatility has been one of the most active areas of research in time series econometrics. Volatility research has not been just limited to the area of time series econometrics dealing with issues of estimation,

[1]See the Bank of England's Annual Report for 2008 published on 14 July 2008 for the role of those two markets in recent volatile events.

statistical inference, and model specification. More fundamentally, volatility research has contributed to the understanding of important issues in financial economics such as portfolio allocation, option pricing, and risk management. Volatility, as a measure of uncertainty, is of most interest to economists and, in particular, to those interested in decision making under uncertainty. Since volatility – the second moment of the distribution of returns – is unobserved, much work has been devoted to measuring, modelling and understanding its evolution. There are several salient features about financial market volatility that are now well documented:

- Fat tails: although the asset returns have different degrees of variation, most of them have fatter tails when compared with the normally distributed random variable. This observation is also referred to as excess kurtosis. The standardized fourth moment for a normal distribution is 3 whereas for many financial series a value well above 3 is observed. Mandelbrot (1963) and Fama (1963, 1965) are the first studies to report this feature.

- Volatility clustering: as noted by Mandelbrot (1963), the observation of large movements is followed by large movements. A quantitative manifestation of this fact is that, while returns themselves are uncorrelated, absolute returns or their squares display a positive, significant and slowly decaying autocorrelation function. Volatility clustering is an indication of persistence in shocks. Correlograms and corresponding Box-Ljung statistics show significant correlations which exist at extended lag lengths.

- Volatility asymmetry: this refers to the idea that price movements are negatively correlated with volatility. This is known as the leverage effect. It was first suggested by Black (1976) for stock returns. Black argued that the fall in stock price causes the leverage and financial risk of the firm to increase. The phenomenon of volatility asymmetry is most marked during large falls. Empirical evidence on leverage effects can be found in Nelson (1991), Gallant, Rossi and Tauchen (1992, 1993), Campbell and Kyle (1993) and Engle and Ng (1993).

- Long memory: as noted, strong volatility persistence, or long memory, is another well-known fact about financial market volatility. It has been extensively discussed (see, e.g., Journal of Econometrics 1996, vol. 73, no. 1). Researchers have noticed that the autocorrelation of the function of returns is slow to decay. High autocorrelation values indicate long memory. Taylor (1986) investigates this interesting phenomenon.

- Co-movements in volatility: the returns and volatility of different assets (e.g. different company shares) and different markets (e.g. stock vs. bond markets in one or more regions) tend to move together. More recent research finds correlation among volatility is stronger than that among returns and both tend to increase during bear markets and financial crises.

Among these properties, the phenomenon of volatility clustering has intrigued many researchers. In detail, volatility clustering refers to the time-varying nature of returns fluctuations, the discovery of which led to Robert Engle's Nobel Prize for his achievement in modelling it. Fluctuations of financial asset returns are 'lumpier' in contrast to the even variations of the normally distributed variable. In the finance literature, this 'lumpiness' is called volatility clustering. With volatility clustering, a turbulent trading day tends to be followed by another turbulent day, while a tranquil period tends to be followed by another tranquil period. Engle (1982) is the first to use the autoregressive conditional heteroscedasticity (ARCH) model to capture this type of volatility persistence: 'autoregressive' because high/low volatility tends to persist, 'conditional' means time-varying or with respect to a point in time, and 'heteroscedasticity' is the technical jargon for non-constant volatility[2]. Volatility clustering has oriented in a major way the development of stochastic models in finance – GARCH (generalized ARCH) models are one of those[3] intended primarily to model this phenomenon.

1.1 Preliminaries and the stylized facts

1.1.1 Prices and returns

A financial market is one in which financial assets can be purchased or sold. One party transfers funds in financial markets by purchasing financial assets previously held by another party. Financial markets facilitate the transfer of funds from surplus units to deficit units. Because funding needs vary among deficit units, various financial markets have been established. The main participants in financial market transactions are households, business (including financial institutions), and governments. Financial markets play a crucial role not only in helping individuals, corporations, and government agencies obtain financing but also in helping individuals or corporations invest in financial assets.

Financial asset prices are dynamic, changing frequently whenever the financial markets are open. A striking feature of financial asset prices is that they move more rapidly during some months than during others. Prices move relatively slowly when conditions are calm, while they move faster when there is more news, uncertainty, and trading. Statistical analysis of market prices is more difficult than analysis of changes in prices.

[2]It is worth noting that the ARCH effect appears in many time series other than financial time series. In fact Engle's (1982) seminal work is illustrated with the UK inflation rate.

[3]GARCH, and stochastic volatility (SV) models are intended primarily to model volatility clustering. They are useful in this pursuit because they are estimated on the basis of return distribution. GARCH models, in addition, are easy to implement.

This is because consecutive prices are highly correlated but consecutive changes have very little correlation, if any. Consequently, it is more convenient to investigate suitable measures of changes in prices. Returns can be defined by changes in the logarithms of prices, with appropriate adjustments for dividend payments. Almost all empirical research analyzes returns to investors rather than prices. Returns are more appropriate for several reasons. The most important is that returns, unlike prices, are only weakly correlated through time.

General properties that are expected to be present in any set of returns are called stylized facts. Three stylized facts of daily returns obtained from a few years of prices are documented and discussed in the literature. First, the distribution of returns is approximately symmetric and has high kurtosis, fat tails and a peaked center compared with the normal distribution. Second, the autocorrelations of returns are all close to zero. Third, the autocorrelations of both absolute returns and squared returns are positive for many lags and they indicate substantially more linear dependence than the autocorrelations of returns.

1.1.2 Volatility

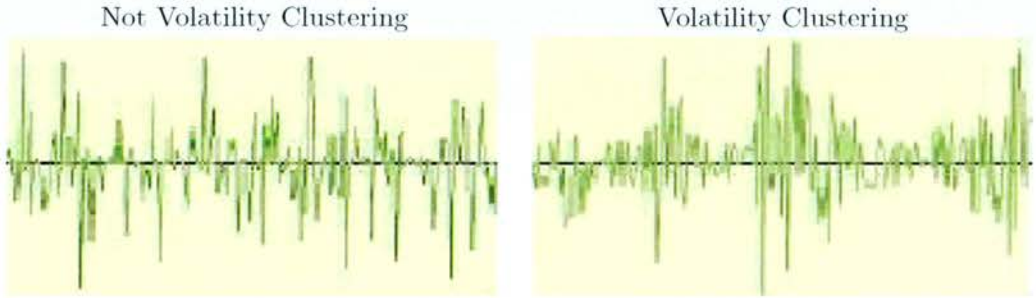
Volatility is a measure of price variability over some period of time. It typically describes the standard deviation of returns. Alternatively, we can say that volatility is the standard deviation of the change in the logarithm of a price or a price index during a stated period of time. Volatility changes explain the major stylized facts for time series of daily returns by assuming that volatility follows a stochastic process, which has the property that today's volatility is positively correlated with the volatility on any future day.

Volatility is a measure for the second moment of a distribution. The first moment is the mean, the third is skewness, and the fourth, kurtosis. So, the first two moments alone are sufficient statistics for summarizing the characteristics of the entire bell-shaped distribution. It is, therefore, convenient to equate return and risk to the first two moments of the return distribution, and indeed, this assumption is fundamental in Markowitz mean–variance portfolio theory and the capital asset pricing model.

A simple measure of volatility considered is referred to as an unconditional measure of volatility because it is made without regard to whether available information is used to predict it. Volatility which is unconditional on the available information set is called unconditional volatility. Realized volatility, also called historical volatility, is the standard deviation of a set of previous returns. Conditional volatility is the standard deviation of a future return that is conditional on known information such as the history of previous returns. The difference between unconditional and conditional volatility is that the term "conditional" implies explicit dependence on a past sequence of observations. The term "unconditional" applies more to long-term behaviour of a time series, and assumes no explicit knowledge of the past.

Volatility clustering, or persistence, suggests a time-series model in which successive disturbances are uncorrelated, yet serially dependent. Stylized facts about volatility clustering include the following: (1) both "good" news (positive shocks) and "bad" news lead to higher levels of volatility (Engle, 1982 and Bollerslev, 1986); (2) bad news tends to increase future volatility more than good news (Nelson, 1991); (3) the effect of news on volatility has a transitory (rapid decay) and more permanent (slow decay) component (Engle and Lee, 1999); and (4) volatility appears to have an effect on the risk premium (Merton, 1980 and French, Schwert, and Stambaugh, 1987).

Note that volatility clustering is a property of most heteroskedastic stochastic processes. Heteroskedasticity is the property of time-varying (conditional or unconditional) variance in a stochastic process. Volatility clustering is the property that there are periods of high and low (conditional or unconditional) variance. The volatility "clusters." This is illustrated in Figure 1.1. Its graph on the left hand side is a realization of a heteroskedastic stochastic process without volatility clustering. Its volatility fluctuates, but it is independent from one time to the next. The graph on the right hand side is a heterokedastic stochastic process with volatility clustering.



Note: Other sources.

Figure 1.1: Volatility clustering

What is the basic insight behind conditional volatility? It is that financial asset prices are partly determined by their volatility. This is embodied in many financial models such as the capital asset pricing model (CAPM) and the Black-Scholes option pricing model. Furthermore, volatility is itself predictable. Indeed, financial market participants frequently make predictions of future volatility on the basis of current and past volatility in order to make decisions about their market behaviour. Traditionally, financial analysts use volatility clustering to predict price volatility. This involves obtaining data on past standard deviations and predicting current and future standard deviations on the basis of some model derived from the past data. The measure of conditional volatility is best seen as an upgraded version of the traditional volatility clustering.

1.2 Motivation

The main reason for the prominent role that volatility plays in financial markets is that volatility is associated with risk and uncertainty, the key attributes in investing, option pricing, and risk management. Heteroscedasticity, a technical term for time-varying volatility, makes the estimation of asset-pricing relationships inefficient. Time series of financial asset returns often exhibit the volatility clustering property. Taking account of these important features on volatility in general and conditional volatility in particular, this thesis concentrates on investigating three important issues throughout the conditional volatility process including forecasting methodology, the theoretical underpinning behind the phenomenon, and optimal forecast evaluation. Throughout this thesis, one of the main strengths of the work is set deeply in the second issue which presents a theoretical underpinning to the stylized fact of conditional volatility in FOREX returns.

Which method will provide the best forecasts?

The foreign exchange (FOREX) market is without any doubt the largest financial market in the world. It has a turnover which exceeds by far other markets such as stocks and bonds.

As exports and imports have grown as a percentage of the gross domestic product (GDP) of all developed countries, so too has the proportion of foreign exchange market participants such as firms and governments earning foreign exchange and/or requiring foreign currencies to purchase intermediate or final goods. Such foreign exchange market participants are necessarily exposed to foreign exchange risk resulting from variations in exchange rates. Foreign exchange market participants have sought both to protect themselves from this risk and to seek profits through speculation on foreign exchange markets. The desire to protect against risk has led to the development of markets designed to provide insurance (forward and futures markets) and the exploitation of techniques such as currency swaps and options. At the same time, more attention has been paid to the need to forecast future changes in exchange rates. The main reason is that if foreign exchange market participants can forecast exchange rate volatility, they can determine the potential range surrounding along with their point estimate forecast and reduce risk for a particular currency.

Exchange rate volatility forecasting is an important input for investment – international investors require portfolio diversification beyond national borders and international traders make export and import decisions, option pricing – options traders require volatility forecasts to price options, risk management – risk managers use internal models such as value-at-risk applications, and financial market regulation – central banks who require interval forecasts whether an exchange rate will fluctuate within a target zone.

Foreign exchange market volatility is clearly forecastable. Research has shown that the one-day-ahead forecasting record for exchange rates is 10–15 percent and is likely to increase by about threefold if the ex post volatility is measured more accurately (Poon and Granger, 2005). A number of exchange rate volatility 'stylized facts' have been documented since the abandonment of the Bretton Woods system of fixed parities more than 36 years ago. First is the phenomenon of volatility clustering where large exchange rate changes are typically followed by other large changes, eventually giving way to more tranquil periods (Baillie and Bollerslev, 1991). And second, periods of high exchange rate volatility have displayed remarkable persistence, in some cases lasting years (Engel and Bollerslev, 1986).

The development of different models for volatility is guided by the stylized facts observed in the data. This leads to a large array of alternative models available to practitioners. When the researcher and/or the practitioner faces so many models, the natural question becomes which one to choose. The current debate in literature focuses on the predictive abilities of several popular models of either the same class or different classes, and how far ahead one can accurately forecast, and to what extent volatility changes can be predicted. In terms of the importance of historical volatility (HIS) and ARCH class models very few papers employ the (almost) full range of them for a specified financial asset together to investigate their predictive abilities. This means they are not able to meet the market's need for econometric techniques in controlling and modeling asymmetric and clustering volatility.

The aim of Chapter 2 is to estimate 73 time series volatility models in an effort to maximize capture of the salient features of exchange rate volatility, and evaluate the models in terms of out-of-sample forecast accuracy and efficiency. Given an objective function, we look for the best predictive ability. What Chapter 2 achieves that has not been done before is a comparison among the almost full range of time series models for forecasting performance, a literature survey, pulling together all the results and summarizing them, plus our own estimations, which also lead to the most successful monthly foreign exchange forecast model by using the forecast evaluation criteria.

Which economic behaviour is consistent with ARCH?

The volatility of FOREX prices is not the same at all times. Volatility clustering is seen in periods of high and low volatility when returns are plotted in time order (Section 1.1.2). Furthermore, the stylized fact that squared returns are positively autocorrelated (Section 1.1.1) is indicative of positive autocorrelation in the volatility process. In later chapters we will see that the parameters of ARCH models reject the hypothesis of constant volatility for the FOREX time series. It is also well known that traders do not believe FOREX volatility is constant, because implied volatilities vary considerably over time.

So why does FOREX volatility change? What economic behaviour explains the origin of volatility clustering in FOREX returns, if any?

While most empirical work in finance acknowledges the volatility clustering in FOREX returns (in particular see Engle et al, 1990), there is little theoretical underpinning to the fact that FOREX returns display autocorrelated volatility characteristics. The purpose of Chapter 3 is to present such a theory.

The modelling of time-varying volatility in the foreign exchange market is largely dominated by models of autoregressive conditional heteroskedasticity (ARCH) and generalized ARCH (GARCH) (see Sarno and Taylor, 2002 for a literature review). Although these models have proved to be useful, they do not explain the economic reasons behind the clustering in FOREX returns. Some of the "accepted" explanations of time-dependent volatility given in the literature are: (1) the amount of information or the quality of information reaching the market is in clusters —or else, the time necessary for market agents to fully process new information— (see Engle et al, 1990 and Baillie and Bollerslev, 1991), (2) the relationship between volatility and trading volume induces volatility to be serially correlated[4] (see among others, Cornell, 1981, Grammatikos and Saunders,1986 and Jorion,1996).

Since recent research has suggested a habit-based explanation to the exchange rate risk premium (see among others Verdelhan, 2006), there is a possibility that habit-based models can also explain clustering in FOREX volatility. We explore this idea and find that a habit persistence version of the theoretical Lucas two-country monetary model, as that proposed in Moore and Roche (2006), is capable of generating predictable

[4]We note, however, that in several studies volume or number of transactions have been used as a proxy for the information arrival process.

conditional volatility in spot FOREX returns. This finding adds to the many interesting asset pricing features that the habit model introduced by Campbell and Cochrane (1999) generates, e.g.: pro-cyclical variations of stock prices, long horizon predictability, counter-cyclical variation of stock market volatility, counter-cyclicality of the Sharpe ratio and the short- and long-run equity premium.

The paper that is closest in its focus to Chapter 3 is McQueen and Vorkink's (2004), which develops a preference-based equilibrium asset pricing model that explains low-frequency conditional volatility in stock returns. Their model allows investors to derive utility from fluctuations in wealth as well as from consumption. The level of financial utility derived from portfolio fluctuations depends on a slow-moving measure of prior investment performance (a mental scorecard) which investors use to measure departures from the habit level of the portfolio. It is the revisions to wealth introduced in the utility function which lead to the ARCH behaviour. Despite the common focus of our research and that of McQueen and Vorkink (2004), i.e. to give a theoretical explanation to volatility clustering, we refer to important differences between both: (1) McQueen and Vorkink (2004) analyze stock returns whereas our Chapter 3 analyses exchange rate returns. (2) McQueen and Vorkink use an extended power utility function that allows investors to derive utility not only from fluctuations in consumption but also from wealth to allow risk premia to change over time. We use instead Campbell and Cochrane's (1999) modified power utility function and do not account for fluctuations in wealth. We rely on this approach because, in economics, we typically think of wealth as an instrument that leads to utility via its ability to buy consumption, rather than the object itself. Further, as Campbell and Cochrane's (1999) preferences generate many interesting asset pricing features, it is both appealing and natural to incorporate their preferences to the problem in hand. (3) McQueen and Vorkink reverse-engineered a model to generate return moments such as clustered volatility and then solved and simulated the model. In other words, the clustering of returns motivated their model. In Chapter 3 we test our theoretical model for clustering itself.

We generate artificial data from the theoretical model using the same parameterization outlined in Moore and Roche (2006), where parameters in the model were mainly taken from direct estimation using US data or from existing literature. We analyze the following models: ARCH, GARCH, GARCH-in-mean (GARCH-M), Exponential GARCH (EGARCH), Threshold GARCH (TARCH), Power ARCH (PARCH) and Component GARCH (CGARCH). Our findings show that there is persistence in volatility in the spot exchange returns, e.g. we have evidence that these models are able to model conditional volatility in spot exchange rates and that asymmetric CGARCH(1,1) and PARCH(1,1) are the best estimating models among them. Most important, using quarterly USD/GBP data for the period 02/1973–10/2005, we find that the volatility autocorrelation of the empirical data displays the same patterns found in our simulations. The impulse response functions (IRFs) of the empirical and simulated ARCH processes are similar, i.e. the theoretical model fits the empirical dynamic behaviour of volatility.

What optimal statistical criterion can lead to utility maximization?

Comparing the forecasting performance of competing models is one of the most important aspects of any forecasting exercise. In contrast to the efforts made in the construction of volatility models and forecasts, little attention has been paid to forecast evaluation in the volatility forecasting literature.

Ideally an evaluation exercise should measure the relative or absolute usefulness of a volatility forecast to investors. However, to do that one needs to know the decision process that will include these forecasts and the costs or benefits that result from using these forecasts. Utility-based criteria, such as that used in West, Edison and Cho (1993), require some assumptions about the shape and property of the utility function. Using these utility-based criteria may well lead to high profits in trading but the problem is that nobody in econometric and statistic analysis uses it. In practice these costs, benefits and utility function are not known and it is usual to simply use measures suggested by econometricians and statisticians.

Popular evaluation measures used in the literature include Mean Error (ME), Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percent Error (MAPE) etc. These are well known criteria (loss functions) and have well known statistical properties. One forecasting method is more accurate than another if its average statistical loss is less. It is found that models classed as accurate due to small statistical loss are not useful in practical situations and may give little guide to the potential profitability, while models that perform poorly on statistical grounds may still yield a profit if used for trading, and vice versa.

The motivation stems from the fact of poor-out-of-sample ARCH forecasting performance when judged on the basis of traditional forecast accuracy criteria versus its good performance when more advanced procedures such as utility-based criteria are employed, whilst these economic evaluation criteria would not be practical if none of their statistical properties is available in econometric or statistic analyses. González-Rivera, Lee and Mishra (2004) find which ARCH model is implied by maximizing utility. Another interesting question is what statistical criterion (based on forecasting errors) does the utility maximization correspond to?

The aim of Chapter 4 is to find an optimal forecast error criterion that would lead investors to select the volatility model that provides maximized economic profitability. In other words, an econometrician or statistician would recognize this utility maximization easily by analyzing its noticeable statistical properties. The optimal forecast error criterion is numerically established with the mapping from errors to wealth under some trading rule, an approximation to the function of forecast error for different levels of risk aversion. The empirical results show that for a higher risk averse investor the optimal forecast error criterion is a weighted average of MAE and MSE but that weights MSE less heavily. The optimality forecast error criterion based on functions of forecast errors for utility maximization under asymmetric loss provides a simple rule for making economic and financial decisions under uncertainty.

It is noted that in Chapter 2 we look at problems associated with volatility models and forecasts and in Chapter 3 at attempts to explain this "clustered" volatility for volatile exchange rates. Whatever the causes are, foreign exchange market participants must try in some way to cope with rapidly changing exchange rates. One response, we saw, is to develop fixed exchange rate system or to take the extra step and move to monetary union. In the absence of fixed exchange rates or monetary union, foreign exchange market participants must take action to protect themselves against that risk. The need for sophisticated risk management in the face of highly volatile exchange rates provides one of the principal reasons for the growth of derivatives markets. These allow foreign exchange market participants to hedge risk by taking out contracts in derivatives markets, which carry the opposite risk to that faced in the underlying markets such as the FOREX markets. In order to make the chapter substantial and interesting, Chapter 4 extends to target options, one of the two principal types of derivatives[5], – the S&P500 stock index options. Options are generally preferable in cases where

[5]The two principal types of derivatives are futures and options. Both are tradable contracts offered by futures markets. Futures promise the delivery of an underlying asset of a specified kind on a given date, although delivery is seldom made. In order to increase tradability, both futures and options are highly standardised. Both offer the possibility of very high rates of profit. Both contracts are offered in relation to exchange rates, short-term and long-term interest rates and stock exchange indices. Both are widely used for speculation as well as for risk management.

a hedger is uncertain about the direction the price of the underlying asset is likely to move, while forward and futures contracts are likely to provide cheaper protection against loss than options, but remove the profit opportunity if prices move in favour of the firm. More importantly, volatility is the most important variable in the pricing of derivative securities, of which trading volume has been quadrupled in recent years. To price an option, we need to know the volatility of the underlying asset from now till the option expires. In fact, the market convention is to list option price in term of volatility units. Nowadays, one can buy derivatives that are written on volatility itself, in which case the definition and measurement of volatility will be clearly specified in the derivative contracts. In these new contracts, volatility now becomes the underlying "asset". So volatility forecast and a second prediction on the volatility of volatility over the defined period will be needed in order to price such derivative contracts. The theoretical principle makes the optimal forecast error criterion proposed in Chapter 4 available for most financial markets including FOREX markets.

1.3 Structure of the thesis

This thesis is divided into five chapters, including this introductory part. The figure below shows the structure of the thesis, giving a road map and linking the chapters intuitively. As seen in Figure 1.2, conditional volatility, the empirical phenomenon of volatility clustering in financial markets captured by ARCH type models, runs through the whole thesis. The three important issues regarding conditional volatility will be investigated and discussed in the three main chapters respectively: (1) forecast modeling; (2) theoretical underpinning; and (3) optimal evaluation, which are the core components throughout the conditional volatility process. Two financial markets are implicated: the foreign exchange market and derivative market.

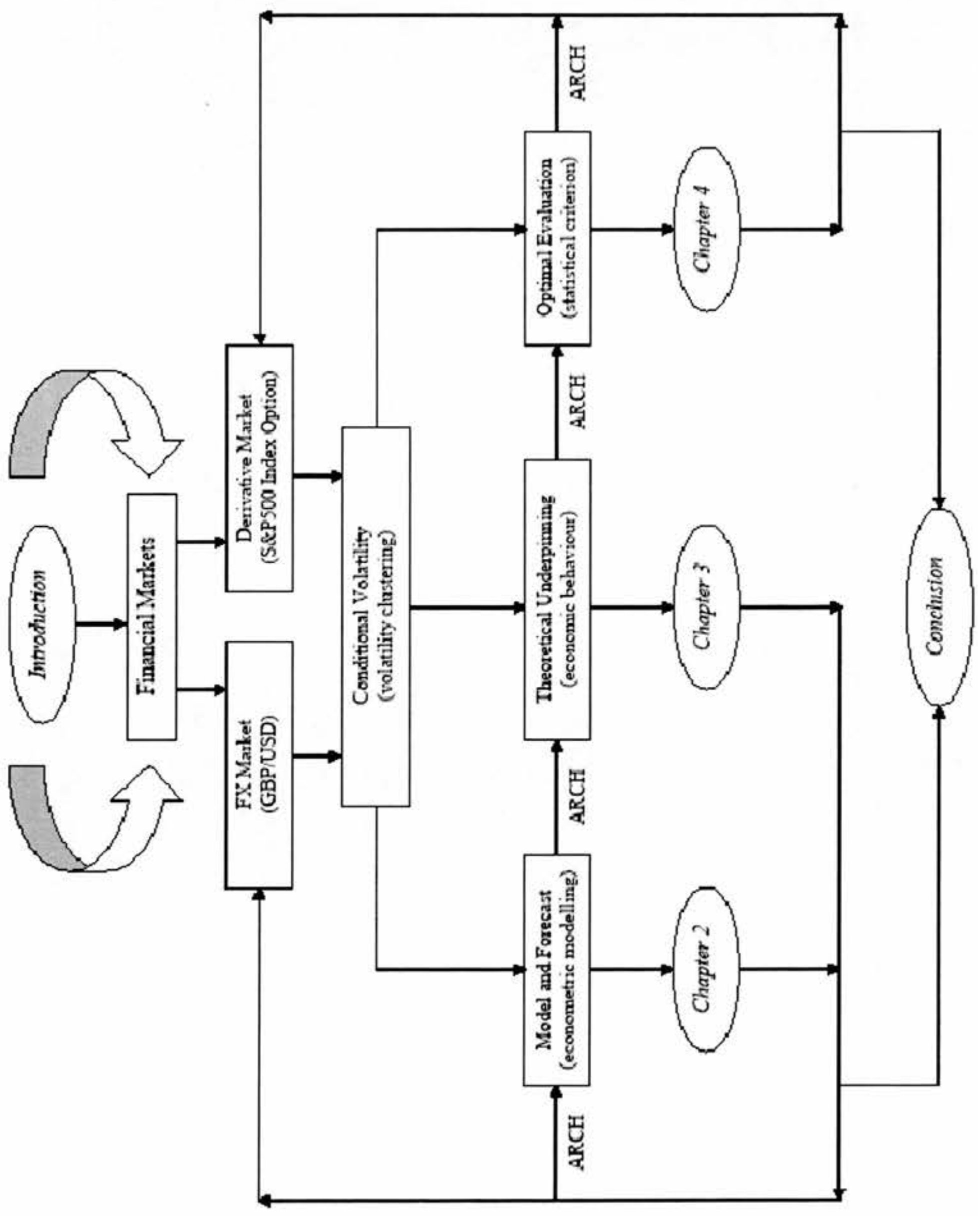


Figure 1.2: Structure of the thesis

The first chapter focuses on what the topic is and why it is important – providing a foundation for the conditional volatility of time series of returns from financial assets, introduces motivations and contributions, and explains the road map of the linkage chapters. Chapter 2 studies volatility forecast modeling in general and the out-of-sample predictive ability of 73 competing time series models for the volatility of foreign exchange changes in particular. Chapter 3 gives a theoretical underpinning to conditional volatility by presenting evidence that a habit persistence version of the theoretical Lucas two-country model is capable of generating predictable conditional volatility in spot foreign exchange returns. Chapter 4 proposes a statistical model selection evaluation criterion of forecast error for conditional forecasts that could lead investors to achieve economic utility maximization in trading S&P500 index options. The last chapter summarizes and concludes what the thesis has achieved, and provides some suggested areas of future research. It is noted that the highlight of this thesis is Chapter 3 using economic behaviour on habit persistence to explain the origin of volatility clustering in FOREX returns.

1.4 Notation

Throughout this thesis, except where otherwise stated, the same symbol may have a different meaning in each chapter due to the limited availability of symbols although most have been used consistently. See notations in each chapter for details.

Part I

Forecasting Volatility

Chapter 2

Forecasting exchange rate volatility

2.1 Introduction

Volatility forecast of exchange rates plays an important role in asset and option pricing, international portfolio diversification, performance measurement, hedging currency risk, risk management, policy making and regulation. To some extent, researchers are interested in the implications of how the prices of exchange rates behave; investors and fund managers can objectively and rationally expect future prices and risks due to their understanding of price behaviour; risk managers scale if the portfolio of exchange rates is risky by measuring and predicting volatility; at the same time, quantitative analysts can drive derivative securities and calculate "attractive" prices; policy-makers can make monetary policy and regulators can manage volatility so that chaos is reduced. Accurate forecasting performance is concerned with competing models, forecast horizon and sample frequency.

A number of exchange rate volatility 'stylized facts' have been documented since the abandonment of the Bretton Woods system of fixed parities more than 36 years ago. First is the phenomenon of volatility clustering where large exchange rate changes are typically followed by other large changes, eventually giving way to more tranquil periods (Baillie and Bollerslev, 1991). And second, periods of high exchange rate volatility have displayed remarkable persistence, in some cases lasting years (Engel and Bollerslev, 1986).

Considering the important role of exchange rates to the market participants making financial investment decisions, and the relative unpopularity of foreign exchange (FOREX) markets compared with stock markets, this chapter aims to test the relative quality of exchange rate volatility forecasts generated by a number of models, including both historical volatility and autoregressive conditional heteroskedasticity (ARCH) type models. A larger data set of the US Dollar / UK Sterling (USD / GBP) exchange rate than previous studies forecasting volatility at the monthly forecast horizon will also be used. Using calendar month as the forecast horizon is uncommon in the previous research. Most of the literature forecasts are at the higher frequency. Monthly (or even quarterly) volatility is more relevant to economic models, such as that of Lucas in Chapter 3, than option pricing although it is useful for options. It is useful to analyze further the relationship between the expected volatility and macroeconomic variables since the latter are often made publicly available in monthly announcements.

The aim of this chapter is to estimate conditional volatility models in an effort to capture the salient features of exchange rate volatility, and evaluate the models in terms of out-of-sample forecast accuracy and efficiency. The chapter carries out a study into the distinct stylized facts, in particular, into the properties of the volatility of FOREX changes, by comparing the out-of-sample predictive ability of the competing time series models in an empirical framework. The out-of-sample forecasting performance of the monthly USD/GBP volatility is compared for the period between the post-Bretton Woods era and the present day. Forecast efficiency and accuracy tests are used as forecast evaluation criterion. The various model ranks are shown to be sensitive to the error statistics used to assess the accuracy of the forecasts. The empirical results suggest that non-ARCH class models are superior to ARCH class models. However, ARCH class models take predominance when over-predictions are more heavily penalized.

In the chapter, we employ time series (historical and ARCH) not option implied volatility models to forecast exchange rate volatility, although it is well documented that the latter beats time series models for forecasting purpose due to its large and timely information set, which covers current and future as well as historical information. The reasons why we do not use the option-based volatility model are listed as follows: (1) option assets and option exchange trading are limited; (2) option-implied volatility is not unique; (3) time-series can capture volatility persistence; (4) time-series models can match and trace the realized volatility; (5) time-series models can detect the volatility of high-frequency data. At the same time, there are three types of the time series volatility forecasting models: historical volatility, ARCH, and stochastic volatility models. We only consider two out of three classes of the time series models – historical volatility and ARCH type models, because (1) volatility structure is limited by return distribution in stochastic volatility models; (2) there is little relevant literature on stochastic volatility models; (3) the superior performance of stochastic volatility models to the other two-class time series models has not been established. Poon and Granger (2003) compare the volatility forecasting performance of various models in pairs using 66 studies. For their comparisons between historical and ARCH class models, they found that historical volatility forecasting models were more accurate than ARCH class models for accurate forecasting performance. Our empirical results are consistent with these findings.

The chapter's main contribution is to maximize capture of the properties of FOREX volatility by using 73 time series models under different window forecast procedures and the longest sample period of the data (from the beginning of floating exchange rates system to very recent period). What the chapter achieves that has not been done before is a literature survey (pulling together all the results and summarizing them) plus my own estimations, which prefer HISVOL models and are also consistent with the previous research, when using the forecast models without theoretical foundations and the forecast evaluation criteria without economic assumptions for the monthly FOREX volatility forecasts.

The outline of the rest of the chapter is organized as follows. Section 2.2 gives a general literature review. The descriptions of the data are presented in Section 2.3 and the methodology of each of the models used for forecasting, as well as how they might be estimated, are given in Section 2.4. Evaluation techniques are displayed in Section 2.5. Section 2.6 shows the empirical results. The conclusion is summarized and a further research direction in future is introduced in Section 2.7.

2.2 Literature review

Forecasts of volatility are important when assessing and managing the risks of portfolios. A remarkable variety of methods have been used and the conclusions obtained often appear to be contradictory. This variety reflects the fact that volatility is inherently unobservable, so that forecasts must be made of related observable quantities. Poon and Granger (2003) provide a comprehensive survey of recent volatility forecasting studies. Most studies only predict the volatility of one asset or portfolio. Alexander (2001) covers the more general problem of predicting variances and covariances, within a multivariate context.

In this section, we review major findings in the papers that construct volatility forecasts based on historical information only. First, we review the major findings of the main financial markets in Section 2.2.1. Second, we narrow our review to those that investigate exchange rate volatility forecasting in Section 2.2.2.

2.2.1 Comparisons of historical forecasts

Numerous comparisons of the accuracy of naive, EWMA, ARCH, and other historical volatility forecasts are discussed in Poon and Granger (2003). Of particular interest are comparisons between multi-parameter methods (such as GARCH) and single-parameter methods (such as EWMA). These comparisons should avoid in-sample parameter optimization, because different conclusions can arise for out-of-sample forecasts (see, for example, Dimson and Marsh, 1990; Ederington and Guan, 2002).

Taylor (1987) is one of the earliest to test time series volatility forecasting models before ARCH/GARCH permeated the volatility literature. Taylor (1987) studies the use of high, low, and closing prices to forecast DM/$ futures volatility and finds a weighted average composite forecast to perform best. Wiggins (1992) also gives support to extreme value volatility estimators.

In the pre-ARCH era, there were many other findings covering a wide range of issues. Dimson and Marsh (1990) find ex ante time-varying optimized weighting schemes do not always work well in out-of-sample forecasts. Still (1993) finds S&P 500 volatility is higher during recession and that commercial T-Bills spread helps to predict stock market volatility. Alford and Boatman (1995) find, from a sample of 6879 stocks, that adjusting historical volatility towards volatility estimates of comparable firms in the same industry and size provides a better five-year ahead volatility forecast. Alford and Boatman (1995), Figlewski (1997), and Figlewski and Green (1999) all stress the importance of having a long enough estimation period to make good volatility forecasts over a long horizon.

The early study of Taylor (1986) makes several comparisons between EWMA and GARCH(1,1) predictions of the next daily absolute return for forty assets, including stocks, commodities, and currencies. Out-of-sample comparisons of mean squared errors marginally favor the EWMA approach when averages are taken across all the series. The recommended values of the smoothing parameter γ are 0.04 for equities and 0.1 for other assets. Both the EWMA and GARCH predictors are more accurate than the prior sample mean for every series. In contrast, Akgiray (1989) finds GARCH is a more accurate predictor than EWMA for monthly realized variances calculated from daily CRSP (the Center for Research in Security Prices) index returns between 1963 and 1986.

There is no consensus about the relative accuracy of historical volatility forecasts for equity markets. Tse (1991) and Tse and Tung (1992) prefer EWMA forecasts, respectively for Japan and Singapore. Brailsford and Faff (1996), however, find EWMA is poor for the Australian market and they favor forecasts from the GJR(1,1) specification. Franses and van Dilk (1996) disagree. They recommend the QGARCH specification for five European markets and find the GJR is much less accurate. Heynen and Kat (1994) evaluate but do not recommend asymmetric specifications; instead they prefer stochastic volatility forecasts to GARCH(1,1) and EGARCH(1,1) forecasts for seven major equity markets. Balaban, Bayar, and Faff (2005) covers fourteen countries. Their most accurate forecasts of weekly and monthly volatility, obtained from daily index returns, are given by exponentially weighted averages. Gonzalez-Rivera, Lee and Mishra (2004) consider comparing the performance of various volatility models on the basis of economic and statistical loss functions. Their study revealed that there does not exist a unique model that can be regarded as the best performer across various loss functions. The variety of conclusions must be a consequence of using a variety of markets, data frequencies, and loss functions. Many of the apparent differences in accuracy across methods may not be statistically significant, as there are often a small number of independent out-of-sample forecast errors.

2.2.2 Exchange rate volatility forecasts

Volatility forecasting of exchange rates is a hot topic in the financial literature. Over a 20-year development period, at the time of writing, there have been at least 80 published and working papers studying the forecasting performance of various volatility models in FOREX markets. Most of them have been developed within the last 12 years since 1996. Given its important role and that so much has been written on volatility forecasting in FOREX markets, the focus here is on the main concerns of the 80 papers and the collective findings in this pool of research. Relevant research and literature on model forecasting performance using daily FOREX data (excluding intraday) and one-step-ahead horizon are mainly reviewed here. There are short summaries of each of the 80 papers in Appendix A.

Exchange rate volatility may be easier to predict. Forecasts from the GARCH(1,1) specification are recommended in the study of five currencies by Heynen and Kat (1994). They consider an out-of-sample period from 1988 to 1992. West and Cho (1995), however, find that a constant is more accurate than GARCH and related forecasts, when making out-of-sample forecasts of the squares of weekly returns from five exchange rates between 1981 to 1989. This negative result may be a consequence of using weekly observations. Taylor (1987) instead uses daily high, low, and close prices, which are used to define a variety of DM/$ volatility forecasts that are more accurate than a constant during a short out-of-sample period from 1982 to 1983.

It is well known that the frequency of the data, the type of information used and the forecast horizon are all important variables in volatility forecasting. As mentioned previously, Taylor (1986) represents one of the earliest studies in ARCH class forecasts. It finds that the issue of volatility stationary is not important when forecast over the short horizon, non-stationary series (e.g. EWMA) has the advantage of having fewer parameter estimates and forecasts respond to variance change fairly quickly. Bera and Higgins (1997) shows a strong preference for GARCH. Klaassen F. (1998) finds regime switching GARCH superior for volatility forecasting performance and GARCH (1, 1) forecasts less stable. Hu and Tsoukalas (1999) favour the EGARCH model. Park (2002) finds that the out-of-sample volatility forecasts of the regime GARCH (RGARCH) model are apparently superior to both the standard GARCH and random walk models. Vilasuso (2002) finds that the fractional integrated GARCH (FIGARCH) volatility model is better equipped to capture the salient features of exchange rate volatility than the more commonly used GARCH and integrated GARCH (IGARCH) models.

Wei and Frankel (1991), Jorion (1995), Jorion (1996), Campa and Chang (1998) and Benavides (2004) find that the option implied models are superior to the historical models in terms of accuracy. Neely (2004) deepens the implied volatility puzzle and explains the conditional bias found in implied volatility. On the other hand, Taylor (1987), Brooks (1997), Figlewski (1997), Figlewski and Green (1999), and McCrae, Lin, Pavlik, Gulati (2002) find that, over short horizons, ARIMA model forecasts are more accurate for series with moving-average terms of order greater than 1. Taylor (1987) claims the best forecasts are a weighted average of present and past high, low and close prices, with adjustments for weekend and holiday effects. The forecasts can be used to value currency options. Brooks (1997) finds that the random walk model provides reasonably accurate forecasts. Figlewski (1997) finds that forecast of volatility of the longest horizon is the most accurate and historical volatility (HIS) using the longest estimation period is the best, except for short rate – GARCH is the worst in the FOREX market. Figlewski and Green (1999) find that HIS works better than ES in FOREX markets. McCrae, Lin, Pavlik, Gulati (2002) find the ARIMA model superior over short horizons. Tambakis and Royen (2002) find that a random walk is better than the GARCH model.

Several historical forecasts are compared by Ederington and Guan (2002) for long daily time series of returns from DM/$ exchange rate, US equities, the S&P500 index, and US interest-rate securities. The clear winner from their comparisons is a linear function of the EWMA calculated from daily absolute returns. This forecast outperforms a similar construction from squared returns and a variety of forecasts defined by ARCH models.

Some of the papers use non-linear models to capture nonlinear relationships in the FOREX data for forecasting. One of the popular algorithms is neural networks (NNs). For example, Brooks (1997) finds that the random walk model is able to produce reasonably accurate forecasts, while there are modest advantages for non-linear models over random walk and autoregressive models and, in particular, parsimonious neural network and GARCH-type models are effective over a range of series and forecast horizons. Hu and Tsoukalas (1999) find the artificial neural network (ANN) model highly effective. Dunis, Laws, and Chauvin (2002) find that recurrent neural network (RNN) models are the best single modelling approach in a short-term trading context, better than neural network regression (NNR), implied, GARCH (1, 1) and two combination models. Model combination, which has the overall best performing approach in terms of forecasting accuracy, fails to improve the RNN-based volatility trading results.

Other types of forecasting models are available as well. Alizadeh, Brandt, and Diebold (1999) use factor volatility models to forecast volatility and favour two-factor models. Tims and Mahieu (2003) find the multivariate stochastic volatility model fits the exchange rate data quite well. Interestingly, Christoffersen and Diebold (2000) find that volatility forecast ability decays quickly with horizon with model-free procedure. Furthermore, Christoffersen and Diebold (2000) find forecast ability decreases rapidly from 1 to 10 days. Fiess and MacDonald (2002) find that High-Low (HiLo) is superior for Close-to-Close volatility (CLCL). Calvet, Fisher and Thompson (2006) favours bivariate Markov-Switching Multifractal (MSM) in that it performs well in- and out-of-sample relative to a standard benchmark, conditional correlation GARCH (CCGARCH). Dunis, Laws, and Chauvin (2000) find that no single model dominates, though stochastic volatility (SV) is consistently worst. Option implied standard deviation (ISD) based on the Black-Scholes (1973) always improves forecast accuracy. They prefer equal weight combined forecast excluding SV. Lopez (2001) shows forecasts from all models are indistinguishable. There is no single forecasting model suitable for all purposes.

After this brief review, it is found that, in order to compare out-of-sample forecasting performance of the competitive volatility models in FOREX markets, the literature either centres on employing two or three of the same class models or using more than two models from different classes. Very few papers employ the full range of historical volatility and ARCH models for a specified financial asset together to investigate the models' predictive abilities. Forecasting volatility is 'notoriously' complicated and difficult. There is no consensus favouring a single model.

2.3 Data

The data presented in the chapter is based on over twenty years of daily spot exchange rates collected from DataStream[1]. The currency USD/GBP is analyzed, sterling as Benchmark. The sample period covers 19 February 1973, the post-Bretton Woods era, to 31 October 2005, the present day[2]. The raw 8531 daily price observations of exchange rates were transformed into a log-returns series[3] of 8530 daily return observations.

$$R_t = \ln(\frac{P_t}{P_{t-1}}) \tag{2.1}$$

where R_t is the logarithmic daily return of representative daily prices (P_t and P_{t-1}) for USD/GBP in period t, which measures the changes in the logarithms of prices.

Regarding daily logarithmic returns, the standard deviation is employed as the proxy of the actual volatility on that month at the monthly horizon as follows:

$$\sigma_t = \sqrt{\frac{\sum_{t=1}^{n}(R_t - \overline{R})^2}{n-1}}, \quad \text{where } \overline{R} = \frac{1}{n}\sum_{t=1}^{n} R_t \tag{2.2}$$

where σ_t is actual or realized monthly volatility, $\overline{R}$ is the mean daily return of the month, n is the number of the actual trading days in that month (from 20 to 23). In total, there are 393 monthly observations of volatility derived from 8530 daily logarithmic return observations. The first 197 monthly observations are employed as in-sample data for estimation purposes from February 1973 to June 1989; and the last 196 months as out-of-sample data for forecast purposes from July 1989 to October 2005. A complement of standard descriptive statistics of the monthly actual volatility of the sample USD/GBP data is displayed in Table 2.1.

[1]THOMSON DATASTREAM 4.0.

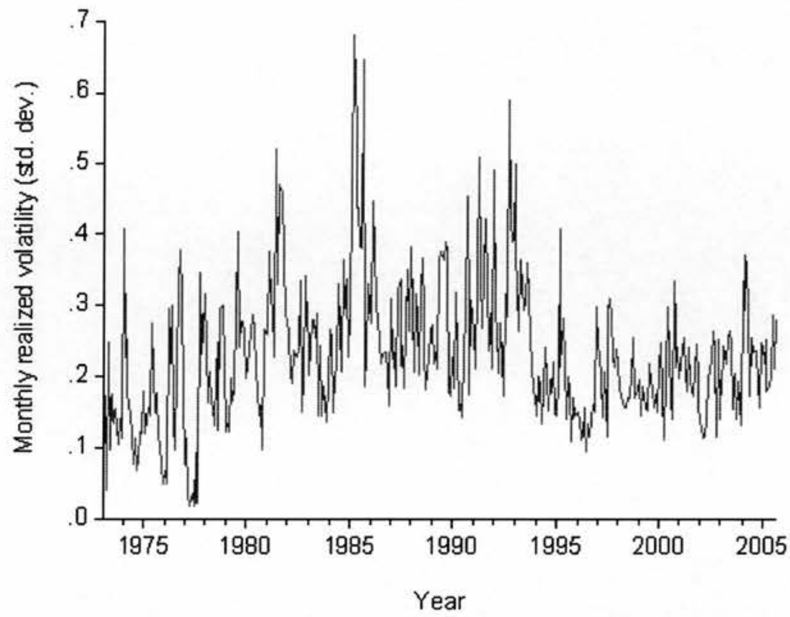
[2]The first draft of the paper was completed in 2006.

[3]Almost all-empirical research analyzes returns to investors rather than prices, Taylor (2005), p9-15.

Vol. Data	Mean	Median	Max.	Min.	Std. Dev.	Skewness	Kurtosis	Jarque-Bera	Probability
USD/GBP	0.231	0.218	0.682	0.017	0.098	0.956	5.082	130.776	0.000

Notes: Std. Dev. (standard deviation) is a measure of dispersion or spread in the series. Skewness is a measure of asymmetry of the distribution of the series around its mean. Kurtosis measures the peakedness or flatness of the distribution of the series. Jarque-Bera test statistic measures the difference of the skewness and kurtosis of the series with those from the normal distribution. The reported probability is the probability that a Jarque-Bera statistic exceeds (in absolute value) the observed value under the null hypothesis of a normal distribution - a small probability value leads to the rejection of the null hypothesis. There are 393 monthly observations.

Table 2.1: Descriptive statistics of monthly realized volatility of daily spot USD/GBP rates for February 1973 - October 2005



Notes: std. dev. denotes standard deviation.

Figure 2.1: Monthly realized volatility of daily spot USD/GBP rates for 1973 - 2005

It is clear from Table 2.1 and Figure 2.1 that the series shows evidence of leptokurtosis. Specifically, the mean and median are 0.231 and 0.218, respectively, while the highest and lowest observations of the series are the 0.682 of March 1985 and 0.017 of April 1977, respectively. The standard deviation is 0.098 measuring the dispersion of the realized volatility series. The distribution has a long right tail with the skewness 0.956 and is leptokurtic relative to the normal because of the kurtosis exceeding 3, 5.082. The Jarque-Bera statistic is 130.78 and the probability is zero so that the null hypothesis of a normal distribution is rejected at the 1% significance level.

2.4 Methodology

Volatility modelling has been a very active area of research in recent years. This interest is largely motivated by the importance of volatility in financial markets. Volatility estimates are widely used as simple risk measures in many asset pricing models. Also volatility enters option pricing formulas derived from models such as the famous Black-Scholes model and its various extensions. For hedging against risk and portfolio management, reliable volatility estimates are crucial.

This very active area of research resulted in the development of several types of models. These alternative models try to account for different stylized facts documented in the literature. In this section, we describe various popular time series volatility models that use the historical information set to formulate volatility forecasts. Brown (1990), Engle (1993), and Aydemir (1998) contain lists of time series models for estimating and modelling volatility. Kroner (1996) explains how volatility forecasts can be created and used. Another important approach is option implied volatility model, which derives market estimates of future volatility from traded option prices. But, as the empirical data is spot rate not option price, the option implied standard deviation based on the Black-Scholes (1973) model and various generalizations will not be discussed here. The implied volatility forecasts are less practical, not being available for all assets, though they use a larger, and more relevant, information set than the alternative methods as they use option prices. Also excluded from discussion here are volatility models that are based on neural networks (Hu and Tsoukalas, 1999), genetic programming (Zumbach, Pictet, and Masutti, 2001), time change and duration (Cho and Frees, 1988; Engle and Russell, 1998).

It is noted that all models described in this section capture volatility persistence or clustering. Others take into account volatility asymmetry also. It is quite easy to construct a supply and demand model for financial assets, with supply a constant and demand partly driven by an external instrument that enters nonlinearity, that will produce a model for financial returns that is heteroscedastic. Such a model is to some extent "theory based" but is not necessarily realistic. The pure time series models discussed in this section are not based on theoretical foundations but are selected to capture the main features of volatility found with actual returns. If successful in this, it is reasonable to expect that they will have some forecasting ability.

2.4.1 Time series volatility forecasting models

2.4.1.1 Predictions based on past standard deviations

Random Walk (RW) Model

This is the simplest stochastic trend model stating conditions that incorporate the wandering ("walking") prices in an unpredictable ("random") manner. Taylor (2005) defines it as follows

$$\widehat{\sigma}_{f,t} = \sigma_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim WN\ (0, \sigma^2) \tag{2.3}$$

where WN denotes a white noise disturbance and $\widehat{\sigma}_{f,t}$, the forecast value at time t, is equal to σ_{t-1}, the realized value at time $t-1$.

Historical Mean (HM) Model

The forecast value $\widehat{\sigma}_{f,t}$ is equal to the historical mean of all of its past observations available before time t, each term with equal weights.

$$\widehat{\sigma}_{f,t} = \frac{1}{t-1}\sum_{i=1}^{t-1}\sigma_{t-i} \tag{2.4}$$

Moving Average (MA) Model

A simple moving average is formed by computing the average (mean) value of the interest over a specified number of previous n periods.

$$\widehat{\sigma}_{f,t} = \frac{1}{n}\sum_{i=1}^{n}\sigma_{t-i} \tag{2.5}$$

In the chapter, $n = 3, 6, 12, 24, 36$, and 60 is employed, respectively.

Weighted Moving Average (WMA) Model

A weighted moving average attaches greater weight to the most recent data. The weighting is calculated from the ratio of the specific time over the sum of previous n periods.

$$\widehat{\sigma}_{f,t} = \frac{n}{N}\sigma_{t-1} + \frac{n-1}{N}\sigma_{t-2} + ... + \frac{1}{N}\sigma_{t-n}, \quad N = 1 + 2 + ... + n \tag{2.6}$$

In the chapter, $n = 3, 6, 12, 24, 36$, and 60 is employed, respectively.

Exponential Smoothing (ES) Model

Exponential smoothing is another modeling technique that uses only the linear combination of the previous values of a series for modeling and generating forecasts of its future values. It converts the observed series, $\{\sigma_t\}_{t=1}^{T}$, into a smoothed series, $\{\overline{\sigma}_t\}_{t=1}^{T}$, and forecast, $\{\widehat{\sigma}_{t+h,T}\}_{t=1}^{T}$, where T is the end of the estimation sample.

$$\begin{aligned} &(a)\ \textit{Initialize at } t = 1\text{: } \overline{\sigma}_1 = \sigma_1 \\ &(b)\ \textit{Update: } \overline{\sigma}_t = \alpha * \sigma_{t-1} + (1-\alpha) * \overline{\sigma}_{t-1}, \quad t = 2, ..., T \\ &(c)\ \textit{Forecast: } \widehat{\sigma}_{f,t} = \overline{\sigma}_t \end{aligned} \tag{2.7}$$

where $0 < \alpha \leqslant 1$ is the damping (or smoothing) factor. The smaller is the α, the smoother is the series $\overline{\sigma}_t$. The h-step-ahead forecasts from single smoothing are constant for all future observations. By repeated substitution, we can rewrite the recursion as $\overline{\sigma}_t = \sum_{i=0}^{t-1} w_i\sigma_{t-i}$ where $w_i = \alpha(1-\alpha)^i$. This shows why this method is called exponential smoothing – the forecast of σ_t is a weighted average of the past values of σ_t, where the weights decline exponentially with time.

Exponential smoothing is a simple method of adaptive forecasting. It is an effective way of forecasting when you only have a few observations on which to base your forecast.

Unlike forecasts from regression models which use fixed coefficients, forecasts from exponential smoothing methods adjust based upon past forecast errors. For additional discussion, see Bowerman and O'Connell (1979).

Exponentially Weighted Moving Average (EWMA) Model

EWMA-n has a built-in mechanism similar to ES for incorporating information from all previous n observations, weighting the information from the closest observations with the highest weight and weights associated with previous observations declining exponentially over time, which means more recent observations will have a stronger impact on the forecast of volatility than older data points. Poon and Granger (2003) describes the EWMA-n model, instead of the observed volatility by the forecast of n-step moving average in ES, as follows:

$$\widehat{\sigma}_{f,t} = \frac{\sum_{i=1}^{n} \alpha^{i} \sigma_{t-i}}{\sum_{i=1}^{n} \alpha^{i}} \tag{2.8}$$

where $0 < \alpha \leqslant 1$ is the damping (or smoothing) factor. EWMA is a truncated version of ES with a finite n. The damping (or decaying) factor could be estimated, but in many studies is set at 0.94 as recommended by RiskMetrics. Two advantages of EWMA over the simple historical model in Brooks (2002) are recent events having more impacts on volatility with more weight than events further in the past and the effect declining exponentially as weights fall. Diebold (2001) note that smoothing techniques in Equations (2.6) and (2.7) produce point forecasts only. They may produce optimal point forecasts for certain special data-generating process, but typically people do not assume that those special data-generating processes are the truth. Instead, the smoothing techniques are used as black boxes to produce point forecasts, with no attempt to exploit the stochastic structure of the data to find a best-fitting model, which could be used to produce interval or density forecasts in addition to point forecasts. In the chapter, $n = 3, 6, 12, 24, 36$, and 60 is employed, respectively.

Regression Model

The regression model is an explicitly multivariate model, where variables are explained and forecasted on the basis of their own history and the histories of other related variables[4].

A conditional autoregression forecasting model is one that can be used to produce forecasts for a variable of interest, conditional upon assumptions about the lagged values of itself, by replacing unknown parameters with estimates. Here, both the autoregression model and autoregression model on standard seasonal and trading-day dummies, where the latter is a 3-variable vector autoregression (VAR) of order 1, are employed. Each of them is assumed to move over time as a first-order autoregression. A first-order autoregression model and its 1-step-ahead point forecast, $\widehat{\sigma}_{f,t}$, at time $t-1$ are following:

$$\sigma_t = \beta_0 + \beta_1\sigma_{t-1} + \varepsilon_t$$

$$\widehat{\sigma}_{f,t} = \widehat{\beta}_0 + \widehat{\beta_1}\sigma_{t-1} \tag{2.9}$$

Furthermore, a first-order autoregression model on both standard seasonal and trading-day dummies and its 1-step-ahead point forecast, $\widehat{\sigma}_{f,t}$, at time $t-1$ are:

$$\sigma_t = \beta_1\sigma_{t-1} + \sum_{i=1}^{s}\gamma_i D_{i,t-1} + \sum_{i=1}^{v}\delta_i^{TD} TDV_{i,t-1} + \varepsilon_t$$

$$\widehat{\sigma}_{f,t} = \widehat{\beta}_1\sigma_{t-1} + \sum_{i=1}^{s}\widehat{\gamma}_i D_{i,t-1} + \sum_{i=1}^{v}\widehat{\delta}_i^{TD} TDV_{i,t-1} \tag{2.10}$$

Where s is seasonal dummy variable, in this chapter, $s = 12$ for monthly volatility forecast. D_i indicates whether we are in the ith month of the year and zero otherwise.

[4]See Diebold (2001), p241, for the details.

For example, D_1 indicates whether we are in January (it is "1" in the first month of the year and "0" otherwise), D_2 indicates whether we are in February (it is "1" in the second month of the year and "0" otherwise), and so on. At any given time $t-1$, we can only be in one of the twelve months, so one seasonal dummy is "1", and all others are "0". $\gamma_i' s$ are seasonal factors which summarize the seasonal pattern over the year. The $TDVs$ are the relevant trading day variables with $v = 3$[5] and its parameters are δ_i^{TD}'s. In no case should we include seasonal dummies and an intercept because of perfect multicollinearity. Alternatively, instead of including a full set of s seasonal dummies, we can include any $s-1$ seasonal dummies and an intercept. This is just a standard regression equation and can be estimated by ordinary least squares.

AR(I)MA Model

ARIMA (autoregressive integrated moving average) models are generalizations of the simple AR model, which use three tools for modelling the serial correlation in the disturbance: (1) the first tool is the autoregressive, or AR, term. An autoregressive model of order p, $AR(p)$ has the form $\sigma_t = \phi_1\sigma_{t-1} + \phi_2\sigma_{t-2} + ... + \phi_p\sigma_{t-p} + \varepsilon_t$; (2) the second tool is the integration order term I; (3) the third tool is the MA, or moving average term. A moving average forecasting model uses lagged values of the forecast error to improve the current forecast. A $MA(q)$ has the form $\sigma_t = \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + ... + \theta_q\varepsilon_{t-q} + \varepsilon_t$. $ARMA(p,q)$ is stated as follows:

$$\widehat{\sigma}_{f,t} = \mu + \phi_1\sigma_{t-1} + \phi_2\sigma_{t-2} + ... + \phi_p\sigma_{t-p} + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + ... + \theta_q\varepsilon_{t-q} + \varepsilon_t \quad (2.11)$$

An $ARMA(p,q)$ model in the variable differenced d times is equivalent to an $ARIMA(p,d,q)$ on the original data, which transforms and makes the original non-stationary data sta-

[5]Initially, we have $v = 5$ trading-day variables to different trading days (8 trading days of the available data in February 1973, 20, 21, 22, and 23 trading days for the rest of other different months) of that different month; finally, $v = 3$ (21, 22, and 23 trading days) is considered in the equation regarding collinearity.

tionary. Box and Jenkins (1976) were the first to approach the task of estimating an *ARMA* model in a systematic manner. Regarding information criteria of model selection, following Diebold (2001), Schwarz's (1978) Bayesian information criterion (SBIC) is employed in the chapter.

In summary, this group of models starts on the basis that past standard deviations are available. In other words, the historical volatilities have to be calculated somehow from historical returns before the volatility model can be estimated. We call this group of models historical volatility models. The various ways of calculating these historical volatilities and the different lengths of sample data used can lead to very different volatility forecasts. The important aspects of using historical models are (1) that actual volatility can be measured accurately and (2) that when high-frequency data are available, that information improves volatility estimation and forecasts.

2.4.1.2 ARCH class conditional volatility models

In the context of financial time series, the variance of the errors is not constant over time, which is known as heteroscedasticity. It makes sense to consider a model which does not assume that the variance is constant, and which describes how the variance of the errors evolves. Another important feature of many series of financial asset returns is known as "volatility clustering". Volatility clustering describes the tendency of large changes in asset prices (of either sign) to follow large changes and small changes (of either sign) to follow small changes. In other words, the current level of volatility tends to be positively correlated with its level during the immediately preceding periods. An immediate question is how could this phenomenon, which is common to many series of financial asset returns, be parameterized (modelled)? One approach is to use an autoregressive conditional heteroskedasticity (*ARCH*) model.

ARCH models are specifically designed to model and forecast conditional variances. The variance of the dependent variable is modelled as a function of the past values of the dependent variable and independent or exogenous variables. They have proved extremely useful for modelling and forecasting volatility fluctuations. In developing an $ARCH$ model, there are three distinct specifications – one for the conditional mean equation, one for the conditional variance and one for the conditional error distribution. $ARCH$ models were introduced by Engle (1982) and generalized as $GARCH$ (Generalized $ARCH$) by Bollerslev (1986) and Taylor (1986). These models are widely used in various branches of econometrics, especially in financial time series analysis. See Bollerslev, Chou, and Kroner (1992) and Bollerslev, Engle, and Nelson (1994) for detailed surveys.

For ARCH-type models several excellent reviews are available including those by Bera and Higgins (1995), Bollerslev, Chou and Kroner (1992), Bollerslev, Engle and Nelson (1994), Diebold and Lopez (1995) and Poon and Granger (2003). These review papers focus on a single class of models; however, this study presents the most popular techniques that are used in the 80 papers reviewed here for modelling exchange rate volatility, and tries to highlight the similarities and differences between them. Due to the space limitations we only cover the issues of interests.

To understand how the model works, a definition of the conditional variance of a random variable, ε_t, is required. The distinction between the conditional and unconditional variances of a random variable is exactly the same as that of the conditional and unconditional mean. The conditional variance of ε_t may be denoted h_t^2, which is written as

$$h_t^2 = var\left(\varepsilon_t \mid \varepsilon_{t-1}, \varepsilon_{t-2}, ...\right) = E\left[\left(\varepsilon_t - E\left(\varepsilon_t\right)\right)^2 \mid \varepsilon_{t-1}, \varepsilon_{t-2}, ...\right] \tag{2.12}$$

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

$$h_t^2 = var(\varepsilon_t \mid \varepsilon_{t-1}, \varepsilon_{t-2}, ...) = E\left[\varepsilon_t^2 \mid \varepsilon_{t-1}, \varepsilon_{t-2}, ...\right] \tag{2.13}$$

Equation (2.13) states that the conditional variance of a zero mean normally distributed random variable ε_t is equal to the conditional expected value of the square of ε_t. Under the ARCH model, the "autocorrelation in volatility" is modelled by allowing the conditional variance of the error term, h_t^2, to depend on the immediately previous value of the squared error

$$h_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 \tag{2.14}$$

The above model is known as an ARCH(1), since the conditional variance depends on only one lagged squared error. Notice that Equation (2.14) is only a partial model, since nothing has been said yet about the conditional mean. Under ARCH, the conditional mean equation describes how the dependent variable, y_t, varies over time. One example of a full model would be

$$y_t = c + \sum_{i=1}^{k} \beta_i x_{it} + \varepsilon_t, \quad \varepsilon_t \sim (0, h_t^2) \tag{2.15}$$

The mean equation given is written as a function of exogenous variables with an error term. Since h_t^2 is the one-period ahead forecast variance based on past information, it is called the conditional variance.

2.4.1.2.1 Symmetric ARCH Models

ARCH Model

$ARCH(p)$'s conditional variance equation is a function of two terms: a constant term, ω, and news about volatility from the previous period, measured as the lag of the squared residual from the mean equation, ε_{t-1}^2, (the $ARCH$ term), which means the expected volatility at time t depends upon the squared forecast errors of the previous periods as follows:

$$h_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 \tag{2.16}$$

GARCH Model

h_t^2 of a GARCH(1,1) model is the one-period ahead forecast variance based on past information

$$h_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}^2 \tag{2.17}$$

The conditional variance equation specified in Equation (2.17) is a function of three terms: a constant term ω and news about volatility from the previous period, measured as the lag of the squared residual from the mean equation ε_{t-1}^2 (the ARCH term), and last period's forecast variance h_{t-1}^2 (the GARCH term).

An ordinary ARCH model is a special case of a GARCH specification in which there are no lagged forecast variances in the conditional variance equation, GARCH(0,1). Higher order GARCH models, denoted $GARCH(p,q)$, can be estimated by choosing either p or q greater than 1 where p is the order of the autoregressive GARCH terms and q is the order of the moving average ARCH terms. The representation of the $GARCH(p,q)$ variance is

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

As stated above, ARCH and GARCH are symmetric conditional volatility models that enforce a symmetric response of volatility to positive and negative shocks.

2.4.1.2.2 Asymmetric ARCH Models

TARCH Model

Threshold ARCH and Threshold GARCH or TARCH, which is also known as the GJR model, were introduced independently by Zakoïan (1994) and Glosten, Jaganathan, and Runkle (1993). The generalized specification for the conditional variance is given by:

$$h_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}^2 + \sum_{k=1}^{r} \boldsymbol{\mu}_k \varepsilon_{t-k}^2 I_{t-k}$$

where $I_{t-k} = 1$ if $\varepsilon_{t-k} < 0$ and $I_{t-k} = 0$ if $\varepsilon_{t-k} \geqslant 0$. In this model, good news, $\varepsilon_{t-k} > 0$, and bad news, $\varepsilon_{t-k} < 0$, have differential effects on the conditional variance; good news has an impact of α_i, while bad news has an impact of $\alpha_i + \boldsymbol{\mu}_k$. If $\boldsymbol{\mu}_k > 0$, bad news increases volatility and there is a leverage effect for the k-th order. If $\boldsymbol{\mu}_k \neq 0$, the news impact is asymmetric. Note that GARCH is a special case of the TARCH model where the threshold term is set to zero. Hence, TARCH(1,1) is

$$h_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}^2 + \boldsymbol{\mu} \varepsilon_{t-1}^2 I_{t-1} \tag{2.18}$$

EGARCH Model

The Exponential GARCH (EGARCH) model was proposed by Nelson (1991). The specification for the conditional variance is

$$\ln(h_t^2) = \omega + \sum_{j=1}^{q} \beta_j \ln(h_{t-j}^2) + \sum_{i=1}^{p} \alpha_i \left[\left| \frac{\varepsilon_{t-i}}{h_{t-i}} \right| - \sqrt{\frac{2}{\pi}} \right] + \sum_{k=1}^{r} \boldsymbol{\eta}_k \frac{\varepsilon_{t-k}}{|h_{t-k}|}$$

Here, the left-hand side of the equation is the log of the conditional variance, which implies that the leverage effect is exponential rather than quadratic, and that forecasts of the conditional variance are guaranteed to be non-negative. The presence of leverage effects can be tested by the hypothesis that $\boldsymbol{\eta}_k < 0$. The impact is asymmetric if $\boldsymbol{\eta}_k \neq 0$. See Knight and Satchell (1998) and Brooks (2002) for the specification of the original elements in detail. Therefore, EGARCH(1,1) is

$$\ln(h_t^2) = \omega + \beta \ln(h_{t-1}^2) + \alpha \left[\left| \frac{\varepsilon_{t-1}}{h_{t-1}} \right| - \sqrt{\frac{2}{\pi}} \right] + \boldsymbol{\eta} \frac{\varepsilon_{t-1}}{|h_{t-1}|} \tag{2.19}$$

PARCH Model

The Power GARCH (PARCH) model is introduced by Taylor (1986), Schwert (1989), and Ding et al. (1993). In the PARCH model, the power parameter θ of the standard deviation can be estimated rather than imposed, and the optional v parameters are added to capture asymmetry of up to order r. The generalized specification for the conditional variance is given by:

$$h_t^{\boldsymbol{\theta}} = \omega + \sum_{j=1}^{q} \beta_j h_{t-j}^{\boldsymbol{\theta}} + \sum_{i=1}^{p} \alpha_i \left(|\varepsilon_{t-i}| - \boldsymbol{v}_i \varepsilon_{t-i} \right)^{\boldsymbol{\theta}}$$

Where $\boldsymbol{\theta} > 0$, $|\boldsymbol{v}_i| \leqslant 1$ for $i = 1, ..., r$, $\boldsymbol{v}_i = 0$ for all $i > r$, and $r \leqslant p$. The symmetric model sets $\boldsymbol{v}_i = 0$ for all i. Note that if $\boldsymbol{\theta} = 2$ and $\boldsymbol{v}_i = 0$ for all i, the PARCH model is simply a standard GARCH specification. The asymmetric effects are present if $\boldsymbol{v}_i \neq 0$. Hence, PARCH(1,1) is

$$h_t^{\boldsymbol{\theta}} = \omega + \beta h_{t-1}^{\boldsymbol{\theta}} + \alpha \left(|\varepsilon_{t-1}| - \boldsymbol{v} \varepsilon_{t-1} \right)^{\boldsymbol{\theta}} \tag{2.20}$$

In the chapter, $\boldsymbol{\theta} = 3$ and $\boldsymbol{\theta} = 4$ are employed.

CGARCH Model

The component GARCH(1,1), $CGARCH(1,1)$, model is described as

$$\begin{aligned} h_t^2 &= q_t + \alpha(\varepsilon_{t-1}^2 - q_{t-1}) + \beta(h_{t-1}^2 - q_{t-1}) \\ q_t &= \omega + \rho\left(q_{t-1} - \omega\right) + \varphi\left(\varepsilon_{t-1}^2 - h_{t-1}^2\right) \end{aligned} \tag{2.21}$$

where q_t represents a time-varying trend or permanent component in volatility, which is driven by the volatility prediction error $(\varepsilon_{t-1}^2 - h_{t-1}^2)$ and is integrated if $\rho = 1$. $h_t^2 - q_t$ describes the transitory component. ρ is typically between 0.99 and 1 so that q_t

approaches ω slowly. An asymmetric component model may be estimated by including a threshold term, which combines the component model with the asymmetric TARCH model, introducing asymmetric effects in the transitory equation and estimates models of the form:

$$\begin{aligned} h_t^2 &= \omega + \rho\left(q_{t-1} - \omega\right) + \varphi\left(\varepsilon_{t-1}^2 - h_{t-1}^2\right) + \alpha\left(\varepsilon_{t-1}^2 - q_{t-1}\right) \\ &\quad + \gamma\left(\varepsilon_{t-1}^2 - q_{t-1}\right) D_{t-1} + \beta\left(h_{t-1}^2 - q_{t-1}\right) \end{aligned} \tag{2.22}$$

where D_{t-1} is the dummy variable indicating negative shocks. $\gamma > 0$ indicates the presence of transitory leverage effects in the conditional variance.

To complete the basic ARCH specification, an assumption about the conditional distribution of the error term ε_t is required. There are three assumptions commonly employed when working with ARCH class models: normal (Gaussian) distribution, Student's t-distribution, and the Generalized Error Distribution (GED). Given a distributional assumption, ARCH models are typically estimated by the method of maximum likelihood. Throughout the dissertation, ARCH models are estimated by the method of maximum likelihood, under the assumption that the errors are conditionally normally distributed.

ARCH Class Models with Dummies

The ARCH class conditional volatility models (symmetric ARCH/GARCH and asymmetric TARCH, EGARCH, PARCH, and CGARCH) combined with the calendar-effect dummies are the models whose conditional variance equation have the dummy variables of calendar effects, which are standard seasonal and trading-day dummies considered as variance regressors. Note that instead of including a full set of s seasonal dummies, we can include any s-1 seasonal dummies and an intercept in the conditional variance equation.

In summary, ARCH-type models define conditional distributions for returns that are characterized by time-varying conditional variances. This characteristic is convenient for the user because daily returns are available for many financial time series. The parameters of these models can be estimated by maximizing the likelihood of observed returns and hence the volatility of returns can be calculated. Many choices can be made in selecting a model, so that an accurate description of the process generating observed returns becomes a realistic aspiration. The disadvantage of such an approach is that the volatility structure is then constrained by the choice of return distribution.

To get reliable forecasts of future volatilities, it is crucial that any satisfactory statistical model for daily returns must be consistent with stylized facts including fat tail distributions of risky asset returns, volatility clustering, asymmetry and mean reversion, and covariation and correlation of volatility across a wide range of assets and financial markets. A number of models employed in the chapter are compatible with these stylized facts. The density of returns is a mixture of normal densities for many of these models, with the mixture defined by variation in volatility. The dependence among absolute and squared returns is then a consequence of slow changes in volatility. As one of the most distinctive stylized facts for most financial time series, volatility clustering and its asymmetry have been well documented. It is emphasized that ARCH models are specifically designed to model and forecast time-varying volatility clustering. ARCH modelling has rapidly become a dominant paradigm when discrete-time models are used to describe the prices of financial assets. It is easy to obtain maximum likelihood estimates of parameters and to compare alternative model specifications. This explains why ARCH models are often preferred to other volatility models that can also explain the stylized facts for returns.

2.4.2 Forecast window and horizon

In the chapter, each of the forecast models mentioned above is run with updating parameters under two different windows respectively: one is a recursive window where the initial estimation date is fixed, but additional observations are added one at a time to the estimation period; the other is a rolling window where the length of the in-sample period used to estimate the model is fixed, so that the start date and end date successively increase by one observation. Static forecast methods are employed, which use the actual rather than forecasted values of lagged dependent variables for one-step-ahead forecasts.

2.5 Forecast evaluation criteria

Good forecasts lead to good decisions. Given the range of techniques available for producing forecasts, it is necessary to select adequate tools for their evaluation. It is popular to focus on absolute evaluation, that is, on testing whether a forecasting model is correctly specified or whether a sequence of forecasts satisfies certain optimality properties. The general idea is to monitor and improve forecasting performance by tracking a record of both forecasts, $\widehat{\sigma}_{f,t}$, at time $t-1$, and realizations, y_t, at time t.

2.5.1 Measures of forecast efficiency

Given optimal forecasting with respect to an information set, the realized volatility series is regressed on the forecasts plus a constant b_0 as follows:

$$\sigma_t = b_0 + b_1\widehat{\sigma}_{f,t} + \epsilon_t \tag{2.23}$$

Perfectly accurate forecasts would imply $b_0 = 0$ and $b_1 = 1$, which is called the "Mincer-Zarnowitz Regression" in Diebold (2001). In this chapter, the 1-step-ahead forecast errors of the optimal forecast are white noise.

2.5.2 Measures of forecast accuracy

The loss function is crucial in measuring forecast accuracy. A few accuracy measures are important and popular. Following Diebld (2001), accuracy measures are usually defined on the forecast errors, $\widehat{e}_{f,t} = \sigma_t - \widehat{\sigma}_{f,t}$, or percent errors, $\widehat{p}_{f,t} = (\sigma_t - \widehat{\sigma}_{f,t})/\sigma_t$. Positive and negative forecast errors can cost differently.

2.5.2.1 Symmetric error statistics

Referring to Diebold et al. (1996), the six main symmetric error statistics employed are mean error (ME), error variance (EV), mean absolute error (MAE), mean absolute percent error (MAPE), mean squared error (MSE) and mean squared percent error (MSPE), where T is the number of out-of-sample observations, as follows:

$$ME = \frac{1}{T}\sum_{t=1}^{T} \widehat{e}_{f,t} \tag{2.24}$$

$$EV = \frac{1}{T}\sum_{t=1}^{T} (\widehat{e}_{f,t} - ME)^2 \tag{2.25}$$

ME measures bias and EV measures dispersion of the forecast errors. The ME and EV are components of accuracy but neither provides an overall accuracy measure.

The most common overall accuracy measures are MSE and MSPE,

$$MSE = \frac{1}{T}\sum_{t=1}^{T} \widehat{e}^2_{f,t} \tag{2.26}$$

$$MSPE = \frac{1}{T}\sum_{t=1}^{T} \widehat{p}^2_{f,t} \tag{2.27}$$

Often the square roots of these measures are used to preserve units, yielding the root mean squared errors (RMSE),

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T} \widehat{e}^2_{f,t}} \tag{2.28}$$

and the root mean squared percent error,

$$RMSPE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\widehat{p}_{f,t}^{2}} \tag{2.29}$$

Somewhat less popular but nevertheless common accuracy measures are the mean absolute error and the mean absolute percent error, respectively,

$$MAE = \frac{1}{T}\sum_{t=1}^{T} |\widehat{e}_{f,t}| \tag{2.30}$$

$$MAPE = \frac{1}{T}\sum_{t=1}^{T} |\widehat{p}_{f,t}| \tag{2.31}$$

2.5.2.2 Asymmetric error statistics

Asymmetric error statistics penalize under-/over-predictions for evaluating volatility with an unequal weight to the same magnitude. Referring to Brailsford and Faff (1996) and Balaban (1999, 2005), mean mixed error (MME) and mean absolute percent error (MAPE) to under-/over-predictions are described

$$MME(U) = \frac{1}{T}\left[\sum_{t=1}^{O} |\widehat{e}_{f,t}| + \sum_{t=1}^{U}\sqrt{|\widehat{e}_{f,t}|}\right] \tag{2.32}$$

$$MME(O) = \frac{1}{T}\left[\sum_{t=1}^{O}\sqrt{|\widehat{e}_{f,t}|} + \sum_{t=1}^{U} |\widehat{e}_{f,t}|\right] \tag{2.33}$$

$$MAPE(U) = \begin{matrix} \frac{1}{T}\sum_{t=1}^{T_U} |\widehat{p}_{f,t}| \text{ if } \widehat{p}_{f,t} > 0 \\ 0 \text{ if } \widehat{p}_{f,t} < 0 \end{matrix} \quad \text{where } T_U \text{ is the number of under-predictions} \tag{2.34}$$

$$MAPE(O) = \begin{matrix} 0 \text{ if } \widehat{p}_{f,t} > 0 \\ \frac{1}{T}\sum_{t=1}^{T_O} |\widehat{p}_{f,t}| \text{ if } \widehat{p}_{f,t} < 0 \end{matrix} \quad \text{where } T_O \text{ is the number of over-predictions} \tag{2.35}$$

where O (U) denotes over-(under-)predictions. MME(U) / MAPE(U) and MME(O) / MAPE(O) penalize the under-predictions and over-predictions more heavily, respectively.

Additionally, the logarithmic error (LE) statistic in Pagan and Schwert (1990) is also asymmetric as follows

$$LE = \frac{1}{T}\sum_{t=1}^{T}\left[\ln(\sigma_t) - \ln(\widehat{\sigma}_{f,t})\right]^2 \tag{2.36}$$

2.6 Empirical analysis

The results of forecast efficiency test are summarized in Tables 2.2a and 2.2b. Regarding perfectly accurate forecasts, models with the non-significant constant term ($b_0 = 0$) and the significant regression parameter term ($b_1 = 1$) against the null hypothesis of zero parameters are efficient. For b_0, we check whether zero is inside the 95% confidence interval for the constant term. If so, the irrelevance is not rejected, which means the null hypothesis is not rejected, otherwise irrelevance is rejected. After the first evaluation, there are 25 models with the zero constant, indicated by the symbol, "^": 21 non-ARCH class models in Table 2.2a and 4 ARCH class conditional volatility models in Table 2.2b. For b_1, the regression parameter is not zero when the irrelevance is rejected, which means checking whether zero is outside the 95% confidence interval for the regression parameter. If so, the irrelevance is rejected otherwise the irrelevance is not rejected. After the second evaluation, there 62 models with the non-zero regression parameter, indicated by the symbol "+" in Table 2.2a and Table 2.2b – 28 non-ARCH class models and 34 ARCH class conditional volatility models.

In order to make sure that the constant and regression parameters are equal to zero and one, respectively, Wald (1) and Wald (2) coefficient tests are used. Wald (1) of the restriction of the regression parameter exhibits the difference between the regression parameter and the unity while Wald (2) of the restrictions, on both the constant and regression parameters, displays the differences of the constant from zero and the

NON-ARCH Models	b_0	t-ratio	Evaluation (1)	b_1	t-ratio	Evaluation (2)	Wald (1)	Wald (2)	Conclusion	$\bar{R}^2$	Rank	SBIC	Rank
RW	0.101	6.74		0.555	8.51	+	*	*		0.310	6	-2.411	6
HM	-0.091	-0.20	^	1.338	0.71					0.002	71	-2.043	71
MA-3	0.075	4.56		0.667	9.11	+	*	*		0.308	7	-2.408	7
MA-6	0.071	3.24		0.687	6.81	+	*	*		0.255	14	-2.335	14
MA-12	0.065	2.08		0.711	4.80	+	***			0.212	17	-2.278	17
MA-24	0.056	1.22	^	0.746	3.47	+			$	0.178	19	-2.236	19
MA-36	0.073	1.54	^	0.670	3.00	+			$	0.123	28	-2.172	28
MA-60	0.084	1.75	^	0.606	2.67	+	***			0.086	37	-2.131	37
WMA-3	0.072	4.75		0.683	10.20	+	*	*		0.337	2	-2.452	2
WMA-6	0.061	3.26		0.731	8.61	+	*	*		0.311	5	-2.412	5
WMA-12	0.050	1.98	^	0.775	6.51	+	***			0.274	12	-2.360	12
WMA-24	0.045	1.18	^	0.797	4.50	+			$	0.222	16	-2.291	16
WMA-36	0.044	0.99	^	0.800	3.91	+			$	0.194	18	-2.256	18
WMA-60	0.049	1.00	^	0.770	3.37	+			$	0.147	24	-2.199	24
ES	0.051	2.88		0.771	9.43	+	*	**		0.337	3	-2.451	3
EWMA_3	0.067	1.81	^	0.700	4.05	+	***			0.176	20	-2.234	20
EWMA_6	0.075	1.81	^	0.664	3.40	+	***			0.148	23	-2.201	23
EWMA_12	0.082	2.44		0.636	3.99	+	**	**		0.168	21	-2.225	21
EWMA_24	0.069	1.45	^	0.691	3.13	+			$	0.152	22	-2.206	22
EWMA_36	0.084	1.77	^	0.619	2.75	+	***			0.105	32	-2.151	32
EWMA_60	0.094	1.95	^	0.565	2.50	+	***			0.075	39	-2.119	39
ARMA(1,1) %	0.001	0.04	^	0.961	9.48	+			$	0.324	4	-2.431	4
ARMA(1,1) @	0.009	0.44	^	0.945	10.27	+			$	0.343	1	-2.461	1
REG @	0.009	0.36	^	0.940	8.42	+			$	0.308	8	-2.408	8
REG %	-0.003	-0.09	^	0.958	7.59	+		***		0.280	11	-2.368	11
REG_D1 @	0.019	0.82	^	0.898	8.65	+			$	0.306	9	-2.406	9
REG_D1 %	0.021	0.75	^	0.861	7.14	+			$	0.261	13	-2.343	13
REG_D2 @	0.033	1.40	^	0.840	8.12	+			$	0.290	10	-2.383	10
REG_D2 %	0.043	1.54	^	0.768	6.34	+	***	**		0.233	15	-2.306	15

Notes: %-Rolling Window; @-Recursive Window; D1-Seasonal dummy variables; D2-Seasonal and Trading-day dummy variables; +-significant; ^-not significant; $-forecast efficiency. At *(1%), **(5%), ***(10%) significant levels for Chi-square statistics; Wald (1) for c(2)=1 test and Wald (2) for c(1)=0, c(2)=1; $\bar{R}^2$-Adjusted R-squared; SBIC-Schwarz Bayesian Information Criterion.

Table 2.2a: Forecast efficiency test for historical volatility models

ARCH Class Models	b_0	t-ratio	Evaluation (1)	b_1	t-ratio	Evaluation (2)	Wald (1)	Wald (2)	Conclusion	$\bar{R}^2$	Rank	SBIC	Rank
ARCH(1) @	0.130	5.54		1.035	3.89	+		*		0.089	36.000	-2.133	36
ARCH(2) @	0.134	5.82		1.021	3.80	+		*		0.091	35.000	-2.135	35
GARCH(1,1) @	0.134	6.36		1.047	4.13	+		*		0.100	33.000	-2.146	33
TARCH(1,1) @	0.132	6.68		1.071	4.61	+		*		0.110	30.000	-2.157	30
EGARCH(1,1) @	0.121	5.84		1.181	4.81	+		*		0.140	26.000	-2.191	26
PARCH(1,1) @	0.087	2.46		0.700	3.82	+		*		0.085	38.000	-2.129	38
CGARCH(1,1) @	0.123	5.58		1.180	4.39	+		*		0.123	27.000	-2.172	27
ARCH(1) %	0.152	5.96		0.821	2.67	+		*		0.052	50.000	-2.094	50
ARCH(2) %	0.163	7.32		0.699	2.56	+		*		0.045	58.000	-2.086	58
GARCH(1,1) %	0.155	6.65		0.834	2.84	+		*		0.065	41.000	-2.108	41
TARCH(1,1) %	0.144	8.12		1.000	5.00	+		*		0.112	29.000	-2.159	29
EGARCH(1,1) %	0.128	5.82		1.161	4.62	+		*		0.142	25.000	-2.193	25
PARCH(1,1) %	0.080	2.17		0.769	3.81	+		*		0.095	34.000	-2.140	34
CGARCH(1,1) %	0.166	6.88		0.705	2.20	+		*		0.048	54.000	-2.089	54
ARCH(1)_D1 @	0.168	8.29		0.637	2.61	+		*		0.059	44.000	-2.101	44
ARCH(2)_D1 @	0.175	8.77		0.573	2.40	+	***	*		0.050	52.000	-2.092	52
GARCH(1,1)_D1 @	0.172	9.17		0.609	2.60	+	***	*		0.055	48.000	-2.097	48
TARCH(1,1)_D1 @	0.197	8.87		0.335	1.15		**	*		0.010	69.000	-2.051	69
EGARCH(1,1)_D1 @	0.168	7.54		0.666	2.24	+		*		0.051	51.000	-2.093	51
PARCH(1,1)_D1* @	0.195	1.42	^	0.066	0.24		*	*		-0.005	72.000	-2.036	72
CGARCH(1,1)_D1 @	0.175	9.28		0.581	2.41	+	***	*		0.050	53.000	-2.092	53
ARCH(1)_D1 %	0.179	9.93		0.533	2.25	+	**	*		0.045	56.000	-2.087	56
ARCH(2)_D1 %	0.182	9.87		0.503	2.11	+	**	*		0.040	61.000	-2.081	61
GARCH(1,1)_D1 %	0.181	10.24		0.519	2.22	+	**	*		0.043	60.000	-2.085	60
TARCH(1,1)_D1 %	0.193	10.34		0.383	1.50		**	*		0.022	66.000	-2.062	66
EGARCH(1,1)_D1 %	0.168	9.82		0.693	2.94	+		*		0.066	40.000	-2.109	40
PARCH(1,1)_D1* %	0.220	1.67	^	0.017	0.06		*	*		-0.005	73.000	-2.035	73
CGARCH(1,1)_D1 %	0.175	8.63		0.612	2.19	+		*		0.044	59.000	-2.085	59
ARCH(1)_D2 @	0.166	8.51		0.659	2.75	+		*		0.063	43.000	-2.106	43
ARCH(2)_D2 @	0.168	7.65		0.648	2.44	+		*		0.056	47.000	-2.098	47
GARCH(1,1)_D2 @	0.138	6.12		1.008	3.43	+		*		0.107	31.000	-2.154	31
TARCH(1,1)_D2 @	0.157	6.11		0.794	2.34	+		*		0.064	42.000	-2.106	42
EGARCH(1,1)_D2 @	0.170	8.20		0.647	2.43	+		*		0.057	46.000	-2.099	46
PARCH(1,1)_D2 @	0.149	3.49		0.288	1.77		*	*		0.013	68.000	-2.054	68
CGARCH(1,1)_D2 @	0.175	8.61		0.589	2.24	+		*		0.046	55.000	-2.088	55
ARCH(1)_D2 %	0.187	9.93		0.444	1.78		**	*		0.030	63.000	-2.071	63
ARCH(2)_D2 %	0.183	8.87		0.488	1.80		***	*		0.033	62.000	-2.074	62
GARCH(1,1)_D2 %	0.169	8.15		0.670	2.31	+		*		0.052	49.000	-2.094	49
TARCH(1,1)_D2 %	0.186	8.07		0.477	1.50			*		0.026	64.000	-2.067	64
EGARCH(1,1)_D2 %	0.175	10.56		0.604	2.78	+	***	*		0.058	45.000	-2.101	45
PARCH(1,1)_D2 %	0.129	2.39		0.346	1.70		*	*		0.014	67.000	-2.055	67
CGARCH(1,1)_D2 %	0.174	8.37		0.618	2.19	+		*		0.045	57.000	-2.087	57
PARCH(1,1)_D2* @	-0.002	-0.01	^	0.439	1.22			*		0.005	70.000	-2.045	70
PARCH(1,1)_D2* %	-0.150	-0.85	^	0.731	2.08	+		*		0.024	65.000	-2.065	65

Notes: %-Rolling Window; @-Recursive Window; D1-Seasonal dummy variables; D2-Seasonal and Trading-day dummy variables; +-significant; ^-not significant; $-forecast efficiency; *-Power Paramete (p)=4 of PARCH otherwise p=3; At *(1%),**(5%),***(10%) significant levels for Chi-square statistics; Wald (1) for c(2)=1 test and Wald (2) for c(1)=0, c(2)=1; $\bar{R}^2$ -Adjusted R-squared; SBIC-Schwarz Bayesian Information Criterion.

Table 2.2b: Forecast efficiency test for ARCH class models regression parameter from unity. Whether the null hypothesis is rejected is based upon the fact that the Chi-square statistic is compared with 1%, 5%, and 10% significant levels, respectively. Furthermore, after Wald (1) and Wald (2) tests, 12 efficient models indicated by the symbol, "\$", are MA-24, MA-36, WMA-24, WMA-36, WMA-60, EWMA-24, ARMA (1, 1) models under both recursive and rolling windows, regression under recursive windows, regressions on seasonal dummies under recursive and rolling windows, and regression on both seasonal and trading-day dummies under recursive window, which are not rejected at 1%, 5%, and 10% significant levels, respectively. It is found that all of the efficient models are non-ARCH class models presented in Table 2.2a. What is more, under each of the adjusted R-squared ($\overline{R}^2$) and SBIC information criterion statistics as goodness-of-fit measure and parsimonious specification of the model, ARMA (1, 1) under recursive window has the best performances with the highest explanatory power $\overline{R}^2$, 34.33%, and the lowest SBIC, -2.461. WMA-3 and ES have the second and third highest explanatory power and lowest information criterion, 33.71% of $\overline{R}^2$ and -2.452 of SBIC for WMA-3 and 33.68% of $\overline{R}^2$ and -2.451 of SBIC for ES, respectively, while PARCH (1, 1) on seasonal dummy with p=4 under rolling window is the poorest fitting models with the negative lowest explanatory power and the highest SBIC.

After the analysis of the forecast efficiency test, the results of the forecast accuracy test are displayed in Tables 2.3a-2.3d. Specifically, symmetric error statistics of forecast accuracy about competitive models are given in Tables 2.3a and 2.3b, where Table 2.3a is for non-ARCH class models and Table 2.3b is for ARCH class conditional volatility models. For components of accuracy, the ME criterion, which is affected easily by outliers, prefers RW most when the absolute values of ME are ranked. Generally, all of non-ARCH class models and six PARCH models on single or double dummies with

Non-ARCH Model	ME	abs ME	Rank	EV	Rank	MAE	Rank	MAPE	Rank	MSE	RMSE	Rank	MSPE	RMSPE	Rank	Symmetry Prob	Rank
RW	-0.001	0.001	1	0.006	24	0.059	16	0.260	12	0.006	0.080	24	0.112	0.335	13	0.033	14
HM	-0.010	0.010	26	0.007	29	0.068	31	0.338	31	0.007	0.085	29	0.193	0.439	29	0.065	31
MA-3	-0.001	0.001	4	0.006	12	0.053	5	0.240	4	0.006	0.075	12	0.104	0.322	5	0.016	5
MA-6	-0.001	0.001	6	0.006	14	0.057	11	0.259	11	0.006	0.076	14	0.110	0.332	11	0.025	11
MA-12	-0.001	0.001	8	0.006	17	0.059	15	0.273	15	0.006	0.077	17	0.123	0.351	15	0.032	13
MA-24	-0.003	0.003	13	0.006	18	0.060	19	0.288	21	0.006	0.078	18	0.141	0.376	22	0.041	19
MA-36	-0.004	0.004	16	0.007	25	0.063	25	0.306	27	0.007	0.081	25	0.163	0.403	27	0.054	26
MA-60	-0.009	0.009	24	0.007	27	0.065	29	0.327	29	0.007	0.083	27	0.204	0.452	30	0.061	29
WMA-3	-0.001	0.001	2	0.005	8	0.052	3	0.235	2	0.005	0.073	8	0.098	0.314	3	0.010	3
WMA-6	-0.001	0.001	3	0.005	7	0.053	4	0.241	5	0.005	0.073	7	0.100	0.316	4	0.011	4
WMA-12	-0.001	0.001	7	0.005	10	0.056	10	0.256	8	0.005	0.074	10	0.107	0.327	7	0.019	7
WMA-24	-0.002	0.002	10	0.006	13	0.058	13	0.273	14	0.006	0.076	13	0.124	0.352	16	0.029	12
WMA-36	-0.003	0.003	14	0.006	16	0.060	18	0.284	18	0.006	0.077	16	0.135	0.368	19	0.037	16
WMA-60	-0.005	0.005	22	0.006	20	0.062	23	0.304	26	0.006	0.079	20	0.162	0.403	26	0.051	24
ES	-0.001	0.001	5	0.005	3	0.051	2	0.235	1	0.005	0.071	3	0.095	0.308	1	0.006	1
EWMA_3	-0.002	0.002	11	0.006	19	0.061	20	0.285	19	0.006	0.079	19	0.134	0.366	18	0.039	18
EWMA_6	-0.002	0.002	12	0.006	23	0.062	24	0.296	25	0.006	0.080	23	0.147	0.383	24	0.049	23
EWMA_12	-0.002	0.002	9	0.006	22	0.062	21	0.288	22	0.006	0.080	22	0.138	0.371	21	0.043	20
EWMA_24	-0.003	0.003	15	0.006	21	0.062	22	0.296	24	0.006	0.080	21	0.148	0.385	25	0.047	22
EWMA_36	-0.004	0.004	18	0.007	26	0.064	27	0.313	28	0.007	0.082	26	0.171	0.414	28	0.057	28
EWMA_60	-0.010	0.010	25	0.007	28	0.066	30	0.331	30	0.007	0.084	28	0.211	0.459	31	0.064	30
ARMA(1,1) %	-0.008	0.008	23	0.005	2	0.053	7	0.257	10	0.005	0.070	2	0.112	0.335	12	0.021	9
ARMA(1,1) @	-0.004	0.004	17	0.005	1	0.051	1	0.238	3	0.005	0.069	1	0.097	0.312	2	0.009	2
REG @	-0.005	0.005	21	0.005	4	0.053	6	0.252	6	0.005	0.071	4	0.106	0.326	6	0.017	6
REG %	-0.013	0.013	29	0.005	9	0.057	12	0.282	17	0.005	0.073	9	0.133	0.365	17	0.034	15
REG_D1 @	-0.005	0.005	19	0.005	5	0.054	8	0.254	7	0.005	0.071	5	0.107	0.328	8	0.019	7
REG_D1 %	-0.013	0.013	27	0.006	11	0.058	14	0.286	20	0.006	0.074	11	0.137	0.370	20	0.038	17
REG_D2 @	-0.005	0.005	20	0.005	6	0.055	9	0.256	9	0.005	0.072	6	0.110	0.331	10	0.022	10
REG_D2 %	-0.013	0.013	28	0.006	15	0.059	17	0.290	23	0.006	0.076	15	0.143	0.378	23	0.045	21

Notes: abs ME-absolute value of ME; %-Rolling Window, @-Recursive Window;D1-Seasonal dummy variables; D2-Seasonal and Trading-day dummy variables;*-Power Parameter p=4 of PARCH otherwise p=3;Giving the rankings for each model and criterion, for forecasts made 1 step ahead, a ranking of '1' indicates the best model, and the same rank denotes two models yielding identical performance to six decimal places.

Table 2.3a: Forecast accuracy test of symmetric error statistics for historical volatility models

ARCH Class Model	ME	abs ME	Rank	EV	Rank	MAE	Rank	MAPE	Rank	MSE	RMSE	Rank	MSPE	RMSPE	Rank	Symmetry Prob	Rank
ARCH(1) @	0.134	0.134	34	0.024	34	0.134	34	0.543	34	0.024	0.156	34	0.323	0.568	32	0.075	32
ARCH(2) @	0.136	0.136	41	0.025	35	0.136	40	0.553	37	0.025	0.158	35	0.333	0.577	35	0.083	36
GARCH(1,1) @	0.138	0.138	51	0.025	44	0.138	49	0.565	47	0.025	0.160	44	0.346	0.588	41	0.102	44
TARCH(1,1) @	0.138	0.138	55	0.025	43	0.138	50	0.566	51	0.025	0.160	43	0.347	0.589	42	0.105	47
EGARCH(1,1) @	0.137	0.137	49	0.025	38	0.137	43	0.566	50	0.025	0.158	38	0.345	0.588	40	0.096	42
PARCH(1,1) @	0.026	0.026	30	0.007	30	0.064	26	0.275	16	0.007	0.086	30	0.116	0.341	14	0.054	27
CGARCH(1,1) @	0.139	0.139	59	0.026	46	0.139	55	0.572	55	0.026	0.160	46	0.352	0.594	45	0.113	51
ARCH(1) %	0.135	0.135	39	0.025	39	0.135	37	0.548	36	0.025	0.158	39	0.330	0.574	33	0.083	36
ARCH(2) %	0.135	0.135	37	0.025	40	0.135	36	0.548	35	0.025	0.159	40	0.332	0.576	34	0.082	35
GARCH(1,1) %	0.140	0.140	61	0.026	59	0.140	60	0.575	59	0.026	0.162	59	0.359	0.599	54	0.130	60
TARCH(1,1) %	0.143	0.143	69	0.027	68	0.143	69	0.595	69	0.027	0.164	68	0.377	0.614	66	0.151	69
EGARCH(1,1) %	0.142	0.142	68	0.026	60	0.142	66	0.589	68	0.026	0.162	60	0.370	0.608	61	0.142	64
PARCH(1,1) %	0.036	0.036	31	0.008	31	0.065	28	0.268	13	0.008	0.089	31	0.108	0.329	9	0.053	25
CGARCH(1,1) %	0.140	0.140	63	0.027	63	0.141	62	0.578	61	0.027	0.163	63	0.364	0.603	59	0.137	61
ARCH(1)_D1 @	0.135	0.135	35	0.025	37	0.135	35	0.553	38	0.025	0.158	37	0.340	0.583	36	0.081	34
ARCH(2)_D1 @	0.135	0.135	38	0.025	42	0.136	39	0.555	39	0.025	0.159	42	0.343	0.585	39	0.088	40
GARCH(1,1)_D1 @	0.136	0.136	43	0.026	45	0.137	42	0.562	42	0.026	0.160	45	0.349	0.591	43	0.096	43
TARCH(1,1)_D1 @	0.136	0.136	45	0.026	57	0.138	48	0.562	43	0.026	0.162	57	0.352	0.593	44	0.109	50
EGARCH(1,1)_D1 @	0.138	0.138	50	0.026	51	0.138	52	0.568	52	0.026	0.161	51	0.356	0.597	50	0.113	51
PARCH(1,1)_D1* @	-0.274	0.274	71	0.083	71	0.275	71	1.488	71	0.083	0.288	71	2.940	1.715	71	0.158	71
CGARCH(1,1)_D1 @	0.137	0.137	47	0.026	49	0.137	44	0.564	46	0.026	0.161	49	0.354	0.595	47	0.104	45
ARCH(1)_D1 %	0.136	0.136	44	0.026	48	0.137	45	0.565	48	0.026	0.160	48	0.356	0.597	49	0.104	45
ARCH(2)_D1 %	0.136	0.136	46	0.026	50	0.138	46	0.564	45	0.026	0.161	50	0.355	0.596	48	0.106	48
GARCH(1,1)_D1 %	0.138	0.138	52	0.026	58	0.139	53	0.571	54	0.026	0.162	58	0.362	0.602	58	0.123	57
TARCH(1,1)_D1 %	0.138	0.138	53	0.027	61	0.140	59	0.574	57	0.027	0.163	61	0.368	0.607	60	0.130	59
EGARCH(1,1)_D1 %	0.141	0.141	65	0.027	64	0.142	64	0.586	65	0.027	0.164	64	0.376	0.613	65	0.143	65
PARCH(1,1)_D1* %	-0.266	0.266	70	0.079	70	0.267	70	1.447	70	0.079	0.280	70	2.790	1.670	70	0.155	70
CGARCH(1,1)_D1 %	0.142	0.142	67	0.027	69	0.143	68	0.587	66	0.027	0.165	69	0.377	0.614	67	0.150	68
ARCH(1)_D2 @	0.135	0.135	36	0.025	36	0.135	38	0.556	40	0.025	0.158	36	0.342	0.585	37	0.083	36
ARCH(2)_D2 @	0.135	0.135	40	0.025	41	0.136	41	0.557	41	0.025	0.159	41	0.342	0.585	38	0.090	41
GARCH(1,1)_D2 @	0.139	0.139	58	0.026	47	0.139	57	0.575	60	0.026	0.160	47	0.359	0.599	53	0.119	54
TARCH(1,1)_D2 @	0.139	0.139	57	0.026	55	0.140	58	0.575	58	0.026	0.161	55	0.359	0.600	55	0.125	58
EGARCH(1,1)_D2 @	0.138	0.138	54	0.026	53	0.139	56	0.573	56	0.026	0.161	53	0.361	0.601	57	0.122	55
PARCH(1,1)_D2 @	-0.045	0.045	32	0.010	32	0.084	32	0.449	32	0.010	0.100	32	0.353	0.594	46	0.076	33
CGARCH(1,1)_D2 @	0.138	0.138	56	0.026	56	0.139	54	0.570	53	0.026	0.162	56	0.360	0.600	56	0.123	56
ARCH(1)_D2 %	0.136	0.136	42	0.026	52	0.138	47	0.564	44	0.026	0.161	52	0.357	0.597	52	0.107	49
ARCH(2)_D2 %	0.137	0.137	48	0.026	54	0.138	51	0.565	49	0.026	0.161	54	0.357	0.597	51	0.114	53
GARCH(1,1)_D2 %	0.140	0.140	62	0.027	62	0.141	61	0.579	62	0.027	0.163	62	0.370	0.609	62	0.137	61
TARCH(1,1)_D2 %	0.140	0.140	60	0.027	65	0.141	63	0.582	63	0.027	0.164	65	0.375	0.612	64	0.141	63
EGARCH(1,1)_D2 %	0.141	0.141	64	0.027	66	0.142	67	0.588	67	0.027	0.164	66	0.380	0.616	68	0.147	67
PARCH(1,1)_D2 %	-0.058	0.058	33	0.011	33	0.089	33	0.493	33	0.011	0.105	33	0.416	0.645	69	0.087	39
CGARCH(1,1)_D2 %	0.141	0.141	66	0.027	67	0.142	65	0.585	64	0.027	0.164	67	0.374	0.611	63	0.145	66
PARCH(1,1)_D2* @	-0.297	0.297	73	0.095	73	0.297	73	1.597	73	0.095	0.309	73	3.319	1.822	73	0.162	73
PARCH(1,1)_D2* %	-0.289	0.289	72	0.091	72	0.290	72	1.559	72	0.091	0.301	72	3.167	1.780	72	0.160	72

Notes: abs ME-absolute value of ME; %-Rolling Window; @-Recursive Window;D1-Seasonal dummy variables; D2-Seasonal and Trading-day dummy variables; *-Power Parameter p=4 of PARCH otherwise p=3;Giving the rankings for each model and criterion, for forecasts made 1 step ahead, a ranking of '1' indicates the best model, and the same rank denotes two models yielding identical performance to six decimal places.

Table 2.3b: Forecast accuracy test of symmetric error statistics for ARCH class models different two power parameters under two windows over-predict the volatility; the rest of the ARCH class conditional volatility models under-predict the volatility. As another statistic of accuracy components, EV criterion favours ARMA (1, 1) model under the recursive window. EV has the same ranking of models with the MSE criterion because ME is so small that the square of the difference between forecast errors and ME is equal to the square of the forecast errors at six decimal places. All criteria of the overall accuracy measure are reported in Tables 2.3a and 2.3b from the third column to the sixth column. For both MAE and MSE criteria, ARMA (1, 1) under the recursive window is best. Under MAE, ES is the second and WMA-3 is the third while ARMA (1, 1) under the rolling window is the second and ES is the third for MSE. Both MAPE and MSPE support ES as the best model though ARMA (1, 1) under the recursive window is ranked the second and WMA-3 is the third for the MSPE criterion, to which there is contrary ranking as the second for WMA-3 and third for ARMA (1, 1) under the recursive window to MAPE. As the square root of MSE and MSPE, RMSE and RMSPE have the concurrent and same ranking with MSE and MSPE, respectively. PARCH (1, 1) on both seasonal and trading-day dummies with the power parameter, p=4, under recursive and rolling windows is the first and second worst, respectively, at each of six symmetric error indicators. Taking into account the overall performance of the symmetric error statistic in the forecast accuracy test of all of 73 models, it is concluded that ES is the best and ARMA (1, 1) under the recursive window ranks second, WMA-3 third and PARCH (1, 1) with p=4 on both dummies under the recursive window comes last.

The ranking of both non-ARCH and ARCH class models on asymmetric error statistics is presented in Tables 2.3c and 2.3d, respectively, where Table 2.3c is for non-ARCH class models and Table 2.3d is for ARCH class conditional volatility models. Each model

											Asymmetry		Total	
Non-ARCH Model	MME(U)	Rank	MME(O)	Rank	MAPE(U)	Rank	MAPE(O)	Rank	LE	Rank	Prob	Rank	Prob	Rank
RW	0.131	20	0.148	46	0.233	35	0.281	37	0.109	20	0.058	23	0.092	15
HM	0.134	23	0.175	67	0.198	17	0.433	64	0.132	31	0.075	44	0.140	36
MA-3	0.120	11	0.138	12	0.197	16	0.272	36	0.092	10	0.031	4	0.047	5
MA-6	0.127	14	0.147	45	0.218	32	0.286	40	0.096	11	0.053	17	0.077	11
MA-12	0.130	17	0.150	49	0.214	30	0.315	47	0.101	15	0.058	23	0.091	14
MA-24	0.132	22	0.153	54	0.210	27	0.345	56	0.107	19	0.066	33	0.107	22
MA-36	0.138	30	0.156	59	0.204	23	0.389	61	0.116	25	0.073	41	0.127	28
MA-60	0.137	28	0.160	62	0.187	10	0.441	65	0.129	28	0.071	37	0.133	30
WMA-3	0.121	12	0.135	4	0.182	8	0.282	38	0.088	6	0.025	1	0.035	2
WMA-6	0.119	10	0.141	23	0.200	21	0.270	34	0.088	5	0.034	5	0.046	4
WMA-12	0.125	13	0.146	44	0.205	24	0.293	42	0.091	8	0.049	11	0.068	9
WMA-24	0.128	15	0.150	48	0.209	26	0.318	49	0.099	13	0.056	19	0.085	12
WMA-36	0.130	19	0.153	53	0.211	29	0.335	52	0.103	16	0.063	27	0.100	17
WMA-60	0.135	27	0.156	56	0.199	19	0.386	60	0.113	24	0.069	34	0.120	25
ES	0.118	9	0.138	14	0.192	13	0.266	33	0.084	2	0.026	2	0.032	1
EWMA_3	0.131	21	0.154	55	0.211	28	0.341	55	0.107	18	0.066	32	0.105	20
EWMA_6	0.134	25	0.157	60	0.221	33	0.349	57	0.112	23	0.073	41	0.122	27
EWMA_12	0.134	24	0.156	57	0.228	34	0.331	50	0.110	21	0.069	34	0.112	23
EWMA_24	0.134	26	0.156	58	0.215	31	0.355	58	0.112	22	0.072	39	0.120	25
EWMA_36	0.140	31	0.157	61	0.208	25	0.398	63	0.120	26	0.076	47	0.133	30
EWMA_60	0.137	29	0.161	63	0.190	12	0.444	66	0.131	29	0.074	43	0.137	34
ARMA(1,1) %	0.112	1	0.151	51	0.170	7	0.313	46	0.088	4	0.040	7	0.061	7
ARMA(1,1) @	0.113	3	0.141	24	0.169	6	0.288	41	0.081	1	0.028	3	0.037	3
REG @	0.114	4	0.149	47	0.186	9	0.293	43	0.087	3	0.039	6	0.057	6
REG %	0.112	2	0.166	65	0.188	11	0.331	51	0.098	12	0.052	16	0.087	13
REG_D1 @	0.116	7	0.151	50	0.201	22	0.286	39	0.088	7	0.046	9	0.066	8
REG_D1 %	0.116	6	0.165	64	0.193	14	0.336	53	0.101	14	0.056	19	0.094	16
REG_D2 @	0.117	8	0.151	52	0.194	15	0.297	44	0.091	9	0.047	10	0.070	10
REG_D2 %	0.116	5	0.166	66	0.198	18	0.339	54	0.106	17	0.059	25	0.104	19

Notes: abs ME-absolute value of ME; %-Rolling Window, @-Recursive Window,D1-Seasonal dummy variables; D2-Seasonal and Trading-day dummy variables;*-Power Parameter p=4 of PARCH otherwise p=3;Giving the rankings for each model and criterion, for forecasts made 1 step ahead, a ranking of '1' indicates the best model, and the same rank denotes two models yielding identical performance to six decimal places.

Table 2.3c: Forecast accuracy test of asymmetric error statistics for historical volatility models performs differently under different asymmetric error statistics. ARMA (1, 1) under a rolling window, the regression under a rolling window, ARMA (1, 1) under a recursive window, the regression model under a recursive window rank first, second, third, and fourth, respectively, which could be proof that ARMA (1, 1) is superior than the regression model under the same window; furthermore the rolling window is better than the recursive window at MME(U) while TARCH (1, 1) under the rolling window has the worst performance. PARCH (1, 1) with p=3 under rolling and recursive windows ranks first and second to the best performance, respectively, and ARCH (1) under the recursive window ranks third while PARCH (1, 1) on both seasonal and trading-day dummies with p=4 under recursive and rolling windows performs best and second worst for the MME(O) criterion. According to the MAPE(U) criterion, when p=4, PARCH (1, 1) on seasonal dummy under recursive and rolling windows has the best and second best performance, respectively, and PARCH (1, 1) models on both seasonal and trading-day dummies under recursive and rolling windows rank third and fourth, respectively, which shows that at the MAPE statistic PARCH (1, 1) on dummies with p=4 take priority when under-prediction is heavily penalized. Specifically, PARCH (1, 1) on the single seasonal dummy is superior to PARCH (1, 1) on both dummies. Furthermore, here, the recursive window seems better than the rolling window; again, TARCH (1, 1) under the rolling window is the worst. EGARCH (1, 1) and TARCH (1, 1) under the rolling window, GARCH (1, 1) under the recursive window, and PARCH (1, 1) on both dummies with p=4 under the recursive window rank first, second, third, and last, respectively for the MAPE(O) statistic, although EGARCH (1, 1) under the rolling window is not available when it under-predicts all and over-prediction is penalized, which means that all of its forecasts from EGARCH (1, 1) model under the rolling window are not greater than the realizations, EGARCH (1, 1) under the

											Asymmetry		Total	
ARCH Class Model	MME(U)	Rank	MME(O)	Rank	MAPE(U)	Rank	MAPE(O)	Rank	LE	Rank	Prob	Rank	Prob	Rank
ARCH(1) @	0.348	38	0.135	3	0.548	38	0.084	6	0.845	35	0.044	8	0.119	24
ARCH(2) @	0.351	49	0.136	7	0.558	41	0.062	4	0.893	38	0.051	15	0.134	33
GARCH(1,1) @	0.354	58	0.138	16	0.570	51	0.049	3	0.942	43	0.063	28	0.165	43
TARCH(1,1) @	0.355	62	0.138	15	0.569	49	0.082	5	0.943	44	0.065	31	0.170	45
EGARCH(1,1) @	0.355	59	0.138	11	0.568	48	0.133	13	0.936	42	0.064	30	0.160	42
PARCH(1,1) @	0.164	32	0.130	2	0.254	36	0.301	45	0.124	27	0.053	17	0.107	21
CGARCH(1,1) @	0.357	66	0.139	19	0.574	57	0.102	8	0.968	45	0.072	39	0.185	49
ARCH(1) %	0.349	43	0.136	6	0.553	40	0.127	10	0.888	37	0.050	13	0.133	30
ARCH(2) %	0.348	39	0.136	5	0.552	39	0.131	11	0.916	40	0.050	12	0.132	29
GARCH(1,1) %	0.357	67	0.141	30	0.580	61	0.109	9	1.017	50	0.080	53	0.211	57
TARCH(1,1) %	0.364	73	0.144	39	0.598	73	0.006	2	1.089	62	0.092	66	0.244	68
EGARCH(1,1) %	0.363	72	0.142	33	0.589	68	N/A	1	1.067	58	0.086	62	0.228	61
PARCH(1,1) %	0.176	33	0.119	1	0.266	37	0.271	35	0.132	30	0.050	13	0.103	18
CGARCH(1,1) %	0.357	65	0.143	35	0.588	66	0.102	7	1.068	59	0.086	62	0.223	60
ARCH(1)_D1 @	0.348	40	0.137	8	0.559	42	0.186	21	0.970	46	0.058	21	0.139	35
ARCH(2)_D1 @	0.349	41	0.138	10	0.564	43	0.138	14	1.010	49	0.058	21	0.147	38
GARCH(1,1)_D1 @	0.351	51	0.139	17	0.568	47	0.181	20	1.018	52	0.069	36	0.165	43
TARCH(1,1)_D1 @	0.350	46	0.141	28	0.567	46	0.317	48	1.048	56	0.083	57	0.192	52
EGARCH(1,1)_D1 @	0.352	53	0.140	22	0.573	56	0.261	32	1.005	48	0.078	51	0.191	51
PARCH(1,1)_D1* @	0.276	35	0.514	71	0.090	1	1.502	71	0.847	36	0.079	52	0.237	65
CGARCH(1,1)_D1 @	0.352	52	0.139	18	0.570	50	0.198	23	1.081	60	0.075	45	0.180	46
ARCH(1)_D1 %	0.351	47	0.140	20	0.572	53	0.250	30	1.064	57	0.077	48	0.181	47
ARCH(2)_D1 %	0.351	48	0.140	21	0.571	52	0.236	27	1.089	61	0.077	49	0.183	48
GARCH(1,1)_D1 %	0.352	55	0.141	31	0.578	60	0.206	26	1.149	67	0.088	65	0.212	59
TARCH(1,1)_D1 %	0.352	54	0.143	37	0.578	59	0.384	59	1.188	73	0.104	71	0.234	64
EGARCH(1,1)_D1 %	0.359	70	0.143	36	0.586	65	0.551	68	1.116	63	0.112	72	0.255	72
PARCH(1,1)_D1* %	0.269	34	0.507	70	0.133	2	1.461	70	0.820	34	0.078	50	0.233	63
CGARCH(1,1)_D1 %	0.359	71	0.144	42	0.593	71	0.172	17	1.168	70	0.100	69	0.251	71
ARCH(1)_D2 @	0.349	42	0.138	9	0.564	44	0.179	19	0.991	47	0.060	26	0.142	37
ARCH(2)_D2 @	0.349	45	0.138	13	0.565	45	0.173	18	1.017	51	0.064	29	0.153	40
GARCH(1,1)_D2 @	0.356	63	0.141	29	0.585	64	0.133	12	1.027	54	0.082	55	0.201	54
TARCH(1,1)_D2 @	0.355	60	0.142	32	0.582	62	0.205	24	1.039	55	0.086	64	0.211	58
EGARCH(1,1)_D2 @	0.353	56	0.142	34	0.583	63	0.193	22	1.019	53	0.084	60	0.206	55
PARCH(1,1)_D2 @	0.129	16	0.226	68	0.199	20	0.536	67	0.184	32	0.075	45	0.151	39
CGARCH(1,1)_D2 @	0.354	57	0.141	25	0.576	58	0.206	25	1.125	64	0.085	61	0.207	56
ARCH(1)_D2 %	0.349	44	0.141	26	0.572	55	0.240	28	1.154	68	0.082	54	0.189	50
ARCH(2)_D2 %	0.351	50	0.141	27	0.572	54	0.260	31	1.125	65	0.084	58	0.198	53
GARCH(1,1)_D2 %	0.356	64	0.143	38	0.588	67	0.158	16	1.156	69	0.094	67	0.231	62
TARCH(1,1)_D2 %	0.355	61	0.145	43	0.593	70	0.242	29	1.173	72	0.102	70	0.243	66
EGARCH(1,1)_D2 %	0.358	68	0.144	41	0.590	69	0.397	62	1.136	66	0.113	73	0.261	73
PARCH(1,1)_D2 %	0.130	18	0.241	69	0.169	5	0.598	69	0.207	33	0.072	38	0.158	41
CGARCH(1,1)_D2 %	0.358	69	0.144	40	0.593	72	0.158	15	1.169	71	0.099	68	0.244	69
PARCH(1,1)_D2* @	0.298	37	0.538	73	0.134	3	1.605	73	0.923	41	0.084	58	0.246	70
PARCH(1,1)_D2* %	0.291	36	0.531	72	0.137	4	1.567	72	0.896	39	0.083	56	0.243	66

Notes: abs ME-absolute value of ME; %-Rolling Window; @-Recursive Window;D1-Seasonal dummy variables; D2-Seasonal and Trading-day dummy variables;*-Power Parameter p=4 of PARCH otherwise p=3;Giving the rankings for each model and criterion, for forecasts made 1 step ahead, a ranking of '1' indicates the best model, and the same rank denotes two models yielding identical performance to six decimal places.

Table 2.3d: Forecast accuracy test of asymmetric error statistics for ARCH class models rolling window performs best. What is more, for the LE criterion, ARMA (1, 1) under the recursive window, ES, regression under the recursive window ranks first, second, and third, respectively, and TARCH (1, 1) on the seasonal dummy under the rolling window is worst. Finally, it can be concluded that WMA-3 is the preferred model, followed by ES and ARMA (1, 1) under the recursive window, while EGARCH (1, 1) on both dummies under the rolling window is the least desirable when all five of the asymmetric error statistics are taken into account together. One characteristic of the PARCH models that should be noted is that the combinations between PARCH models and different dummies clearly show different rankings of performances on the asymmetric error statistics, especially for MME(O) and MAPE(U) at which points their superiority is displayed. The inferiority of TARCH (1, 1) under the rolling window should be expected at MME and MAPE, when under-prediction is penalized including its combination of the seasonal dummy at the LE statistic.

In the Tables 2.3a-2.3d, it is also found that ARCH (1) performs relatively better than ARCH (2) selected by SBIC though the latter locates in dominant positions on some symmetric and asymmetric error statistics including (under the rolling window) ARCH (2) for the 35th and ARCH (1) for the 36th on the symmetric error statistics; ARCH (2) for the 12th and ARCH (1) for the 13th on the asymmetric error statistics; and ARCH (2) for the 29th and ARCH (1) for the 30th on the overall symmetric and asymmetric statistics. Furthermore, there is a trend shown that the models without any dummy variable perform better on their own than they do combined with dummies, though dummy models have more additional information. This mirrors the empirical results found by Balaban (1999), with the exception of the PARCH seasonal dummy models at MAPE(U). What is more, for the dummy models, the models on the single seasonal dummy are superior to those on both the seasonal and trading-day dummies, where the effects of seasonality are significant and the trading-day effects are not significant. Finally, the models perform better under recursive windows than they do under rolling windows, meaning the recursive window is preferred. The last column of the Tables 2.3c and 2.3d displays the overall ranking of the competitive models on the overall forecast accuracy evaluation criterion of symmetric and asymmetric error indicators. ES, is the best, and is recommended; WMA-3, the second best and ARMA (1, 1) under the recursive window, third best, although the various model ranks are shown to be sensitive to the error statistics used to assess the accuracy of the forecasts.

Following the super performance of the ES model in terms of forecast accuracy, the further characteristics of ES class models including EWMA-n models with a similar theoretical system to ES are diagnosed in Table 2.4. Specifically, the optimal parameters of ES, EWMA-3 and EWMA-6 are 0.36, 0.13 and 0.12, respectively, where the optimal parameters of EWMA-3 and EWMA-6 chime in with Bowerman and O'Connell's (1979) suggestion that values of optimal smoothing parameter around 0.01 to 0.30 work quite well. 0.99, the optimal estimated damping parameters of EWMA-12, 24, 36, and 60, respectively, are close to one, which is a sign that these series are close to a random walk, where the most recent value is the best estimate of future values. Although 0.94, the damping factor recommended by RiskMetrics, is situated in the range of optimal smoothing parameters, it is not the best for the models of EWMA-6, 12, 24, 36, and 60, respectively. Generally, 0.52-0.99 is the common optimal range of smoothing parameters for EWMA-n models except EWMA-3. The available optimal smoothing parameter range among ES and EWMA-n models, except for EWMA-3 and EWMA-12, is 0.36-0.71 because EWMA-3 has a relatively smaller maximum value (0.46) and EWMA-12 has a relatively greater minimum value (0.52) so that there is no overlapping area between them. It should be emphasized that, for ES models, the codes made in

Model	*Min MSE*	*Optimal Smoothing Parameter*	*Mean of MSE*	*Optimal Range of Smoothing Parameter*	*RiskMetrics*
ES	0.005	0.36	0.005	0.11 - 0.71	/
EWMA-3	0.006	0.13	0.006	0.05 - 0.46	0.94 (out)
EWMA-6	0.006	0.12	0.006	0.07 - 0.99	0.94 (in)
EWMA-12	0.006	0.99	0.007	0.52 - 0.99	0.94 (in)
EWMA-24	0.006	0.99	0.007	0.34 - 0.99	0.94 (in)
EWMA-36	0.007	0.99	0.007	0.35 - 0.99	0.94 (in)
EWMA-60	0.007	0.99	0.007	0.36 - 0.99	0.94 (in)

Notes: The optimal smoothing parameter is α in Equation (2.7) of ES and Equation (2.8) of EWMA-n; We estimate the parameters by minimizing the sum of squared errors; Min MSE-the Minimum Mean Squared Error corresponding to the optimal smoothing parameter; Mean-the mean of the series of all forecasted MSEs with updating parameters by increasing 0.01 at a time from 0.01 to 0.99; the smoothing parameters whose MSEs are less than Mean of MSEs making up Optimal Range of Smoothing Parameter; in and out -inside and outside of Optimal Range of smoothing parameters.

Table 2.4: Optimal smoothing parameters and optimal range of the smoothing parameters for ES and EWMA models

Model	Symmetric Error Statistics												Evaluation	
	ME	*Rank*	*EV*	*Rank*	*MAE*	*Rank*	*MAPE*	*Rank*	*MSE/RMSE*	*Rank*	*MSPE/RMSPE*	*Rank*	*Prob*	*Rank*
Non-ARCH Class Models	15.000	1	15.000	1	15.241	1	16.103	1	15.000	1	16.310	1	15.28	1
ARCH Class Models	51.500	2	51.500	2	51.341	2	50.773	2	51.500	2	50.636	2	51.16	2

Model	Asymmetric Error Statistics										Evaluation		Conclusion	
	MME(U)	*Rank*	*MME(O)*	*Rank*	*MAPE(U)*	*Rank*	*MAPE(O)*	*Rank*	*LE*	*Rank*	*Prob*	*Rank*	*Prob*	*Rank*
Non-ARCH Class Models	15.931	1	48.586	2	20.517	1	49.448	2	15.138	1	22.48	1	16.28	1
ARCH Class Models	50.886	2	29.364	1	47.864	2	28.795	1	51.409	2	46.32	2	50.52	2

Notes: Each of the rank numbers in the table denotes the average ranking of the non-ARCH or ARCH class models on each evaluation criterion, respectively. For the ranking of each class model and criterion, the forecast is made by one step ahead, where a ranking of "1" indicates the best model.

Table 2.5: Summary of forecast accuracy evaluation

the software packages estimate the parameters by minimizing the sum of squared errors, while forecasts from exponential smoothing methods can be adjusted basing upon past forecast errors.

The comparisons between non-ARCH and ARCH class models are summarized in Table 2.5. It is obvious that non-ARCH class models have absolute superiority over ARCH class conditional volatility models on all of the symmetric error criterion for forecast accuracy. For asymmetric error evaluators, although ARCH class conditional volatility models perform better on two asymmetric error statistics of over-prediction: MME(O) and MAPE(O), non-ARCH class models are superior on the other three asymmetric error statistics and overall asymmetric evaluation. Therefore, it can be concluded that non-ARCH class models are preferable to ARCH class conditional volatility models when it comes to forecast accuracy evaluation.

Finally, taking into account the analysis of forecast efficiency and forecast accuracy tests in Tables 2.2a and 2.2b, Tables 2.3a – 2.3d, and Table 2.5, ARMA (1, 1) under the recursive window is recommended because ES and WMA-3 are inefficient although they have the better ranking on forecast accuracy.

2.7 Conclusion

With the move to a flexible exchange rate system in 1973, nominal exchange rate volatility has exhibited remarkable persistence. In this chapter, we compare the out-of-sample forecasting performance of the monthly USD/GBP volatility using the daily foreign exchange prices from 19th February 1973 to 31st October 2005. Non-ARCH (RW, HM, MA, WMA, ES, EWMA, AR(I)MA, regressions, and regressions combined with calendar dummy variables) and ARCH class conditional volatility models (ARCH, GARCH, TARCH, EGARCH, PARCH, CGARCH and ARCH class models combined with calendar-effect dummies) are employed, a total of 73 time series models. The forecasting performance is evaluated by forecast efficiency and accuracy tests. The empirical results favour ES not only on symmetric evaluation but also on overall accuracy, and WMA-3 on asymmetric error statistics, although the various model ranks are shown to be sensitive to the error statistics used to assess the accuracy of the forecasts. Finally, taking account of both forecast accuracy and efficiency, ARMA (1, 1) under a recursive window is recommended. Non-ARCH class models are preferred to ARCH class models though ARCH class models predominate on MME and MAPE of over-prediction that are more heavily penalized. Complex models do not seem to provide better forecasts than simple models. The PARCH model shows its superiority on MME(O) and MAPE(U). A recursive window seems to make more contribution to accurate volatility forecast than a rolling window. For the future, an extension of empirical results to other financial and commodities returns may prove useful.

The volatility forecasting literature is still very active. Many more new results are expected in the near future. There are two areas where future research could seek to make improvements. First is the issue about the sources of volatility changes. Closely related to this is the need to understand the source of volatility persistence. A mere plot of any measure of volatility against time will show the familiar 'volatility clustering' which indicates some degree of forecastability. The biggest challenge lies in predicting changes in volatility. To achieve this we need to understand better the cause of volatility clustering – particularly the economic factors behind the observed volatility clustering in returns. Such an understanding will help to improve time series methods, which are the only viable methods when options, or market-based forecast, are not available. Such a line of research has not yet been pursued vigorously. Second is the issue about forecast evaluation. For example, if statistical tests were conducted to test whether the forecast errors from Model A are significantly smaller, in some sense, than those from Model B, and so on for all pairs. Even if Model A is found to be better than all the other models, the conclusion is not that one should henceforth forecast volatility with Model A and ignore the other models, as it is very likely that models classed as accurate due to small statistical loss are not useful in practical situations and may give little guide to the potential profitability, while models that perform poorly on statistical grounds may still yield a profit if used for trading, and vice versa. It would be useful to find an optimal forecast error criterion that is a linear combination of statistical loss functions. The optimal forecast error criterion would lead investors to select the volatility model that provides maximized economic profitability, while an econometrician or statistician would recognize this utility maximization easily by analyzing its noticeable statistical properties. To find the weights one can run a regression of empirical (expected) utility values on the corresponding forecast errors for different levels of risk aversion. Testing the effectiveness of a composite forecast evaluation criterion is just as important as testing the superiority of the individual models, but this has not been done very often or across different data sets.

Part II

Economic Behaviour behind Conditional Volatility

Chapter 3

Can a Lucas model with habit generate realistic conditional volatility in exchange rate returns?

3.1 Introduction

Time series plots of returns in Figure 1.1 in Chapter 1 display an important feature that is usually called volatility clustering. Volatility clustering describes the general tendency for financial markets to have some periods of high volatility and other periods of low volatility. High volatility produces more dispersion in returns than low volatility, so that returns are more spread out when volatility is higher. A high volatility cluster will contain several large positive returns and several large negative returns, but there will be few, if any, large returns in a low volatility cluster.

In terms of the boom of volatility literatures published in this decade, although much is known about the structure of volatility persistence, little is known about its causes (LeBaron, 2005). People have tried to explain volatility persistence using empirical findings in finance including rate of information arrivals, fashion of price analysis, released economic data, sensitivity of traders to information etc. However, they have not been satisfied with the explanations to volatility persistence, though they have tried to answer questions about: "which economic model or behaviour is consistent with ARCH[1]".

Theoretical asset pricing models, as acknowledged, can explain some variation in volatility while volatility changes explain some stylized facts for asset returns e.g. risk premium. There are some papers that do mention persistent volatility using theoretical asset pricing models for example Campbell and Cochrane (1999) and Moore and Roche (2006). In Campbell and Cochrane (1999) a consumption-based asset pricing model with external habits is able to explain dynamic behaviour of stock prices and even persistent volatility in stock returns. The habit defined by Campbell and Cochrane (1999) as an AR(1) process, in which the lagged level of consumption is the "shock", is a solution to the equity premium puzzle. Moore and Roche (2006) have done pioneering work using the flexible-price two-country monetary model of Lucas's (1982) representative agent theory with habit persistence to solve many FOREX puzzles[2] simultaneously and to mimic the volatilities of real and nominate exchange rates, the forward premium, expected spot returns, and expected forward profits. In Moore and Roche (2006) the utility function depends on surplus consumption. The log of the

[1]We refer "ARCH" afterwards for short to generic predictable conditional volatility.

[2]Moore and Roche (2006) explain the exchange rates disconnect, forward bias, and Meese-Rogoff puzzles in details. With external consumption, inseparable utility and habit persistence, Moore and Roche think that the non-stationary surplus consumption is the radical reason for exchange rate disconnect puzzle; taking an account of the negative correlation between interest rate and expected exchange rate in the nation, Moore and Roche consider that preference for savings triggers forward bias problem.

surplus consumption ratio, consumption and money growths following an AR(1) process is the solution to FOREX puzzles and is able to mimic volatility. As the habit is defined to goods not countries, Moore and Roche (2006) conclude that the Lucas two-country, two-good, two-money economy model with habit is capable of capturing and accounting for FOREX puzzles and some empirical stylized facts in FOREX markets e.g. persistent volatility. Moreover, Moore and Roche (2002, 2004, 2006) note that the surplus consumption ratio is very volatile in comparison to nominal fundamentals such as consumption and money. We use the moment expressions in Moore and Roche (2006) to provide economic insights into simulated results. We agree with Moore and Roche that "the volatility of the fundamentals is able to produce the volatility in the nominal exchange rate". But, the volatility mentioned in both Campbell and Cochrane (1999) and Moore and Roche (2006) is unconditional. Campbell and Cochrane (1999) and Moore and Roche (2006) do not say anything about conditional volatility, an important empirical fact in finance, while, most empirical work involves volatility clustering in returns over the last decade.

Volatility is conditional and asymmetric. There has been little work done investigating the ability of the theoretical asset pricing model with habit to generate volatility clustering in asset returns. McQueen and Vorkink's (2004) study is the one of few examples that applies the theoretical models to the issue of volatility clustering. In McQueen and Vorkink (2004), a preference-based equilibrium asset pricing model is developed to capture long-term stock predictability and excess volatility. The optimal proceeds are made from both consumption and financial utility. McQueen and Vorkink let utility depend on consumption plus the score coefficient multiplied by changes in wealth. They make the marginal utility of financial wealth an AR(1) process by adding a wealth term with a time varying score coefficient. McQueen and Vorkink (2004) show that the mental scorecard that records the market's sensitivity to news and affects the agents' level of risk aversion due to wealth changes and experience loss aversion is able to explain the conditional volatility and even its asymmetric property. But McQueen and Vorkink focus on asset returns rather than exchange rate changes. They do not give the details of the properties of volatility clustering produced by the preference-based equilibrium asset pricing model. For example, how persistent and asymmetric conditional volatility is? What is the dynamic form conditional heteroscedasticity takes? Do its dynamics match those found in the actual data? McQueen and Vorkink (2004) refute the capability of the theoretical model, where utility depends on surplus consumption to explain volatility clustering in U.S. stock data. A clear and thorough literature, which intends to investigate if the consumption-based equilibrium asset pricing model with habit can generate conditional volatility in FOREX returns, is missing.

This chapter's purpose is to investigate volatility clustering in FOREX returns. The chapter develops upon the theoretical model in Moore and Roche (2006) in order to fully investigate the idea of conditional volatility mentioned by McQueen and Vorkink (2004). The chapter works with both the theoretical and empirical frameworks. In the theoretical framework, we use artificial data as in Campbell and Cochrane (1999), McQueen and Vorkink (2004) and Moore and Roche (2006)[3]. We first overview the development of the theoretical model by introducing and discussing several important academic papers, and then deduce an implied ARMA(2,2) process for spot return from the model. Furthermore, we numerically solve the model and find ARCH effects in the spot return we simulate, where the simulated data for spot return and for its innovation in the spot return process are definitely conditionally heteroscedastic. What is more,

[3]Campbell and Cochrane (1999) and McQueen & Vorkink (2004) simulate stock data at a monthly frequency while Moore and Roche (2006) simulates FOREX data at a quarterly frequency with the different parameter settings.

we estimate and establish the form of the conditional heteroscedasticity implied by the model. We find the same fit GARCH models for the best estimates for the simulated and empirical data, which is consistent with the results forecasted in Chapter 2. The estimates of conditional volatility for the innovation in the theoretical quarterly spot return of an ARMA(2,2) process are highly consistent with those for the real monthly spot return itself. We explain why empirical researchers tend to consider the FOREX spot return itself rather than its innovation, which is the reason why we estimate and forecast conditional volatility for the spot FOREX returns in Chapter 2. We show that the dynamics of the conditional heteroscedasticity implied by the model match those we found from the empirical data in Chapter 2 due to the "same looks" of the two impulse response functions (IRFs) for the theoretical and empirical ARCH processes we estimate. The Lucas model with habit can generate realistic conditional volatility in FOREX returns.

The chapter is organized as follows. In Section 3.2 we give an overview of the classic CCAPM model. In Section 3.3 we extend CCAPM to habit persistence using Campbell and Cochrane (1999). In Section 3.4 we review McQueen and and Vorkink (2004) who mobilize the preference-based equilibrium asset pricing model to explain volatility clustering by revisions to wealth introduced in the utility function. In Section 3.5 we present Moore and Roche's habit version of Lucas and show the process in which the model can generate intrinsic conditional volatility in spot returns. In Section 3.6 we numerically solve the model and test the spot return we simulate for the implied time series properties and conditional heteroscedasticity, as well as assess sensitivity to parameter changes. In Section 3.7 we estimate the best fit GARCH model(s) and present IRFs to establish the exact dynamic form of conditional heteroscedasticity. Section 3.8 summarizes and concludes.

3.2 Review of the standard CCAPM

An investor must make a decision how much to consume and how much to save, and what portfolio of assets to hold. The basic idea of most pricing equations is to take the first-order condition (FOC) for that decision. The investor should always set the marginal utility loss of less consumption and more investment today equal to the marginal utility gain of more consumption of the asset's payoff tomorrow. The stochastic discount factor is equivalent to the discounted ratio of marginal utilities. The asset's price should equal the expected discounted value of the asset's payoff. Interests rates are related to the expected trend of consumption because interests rates are related to expected marginal utility. High real interest rates induce more saving and purchase of bonds today and then more consumption tomorrow. High real interest rates should be associated with an expectation of consumption growth. There is an approximately direct proportional relationship between the consumption growth and the interest rate while an approximately inversely proportional relationship between the stochastic discount factor and the interest rate.

As a fundamental measure, marginal utility works for investment performance. The main theory of asset pricing is about how to use marginal utility to solve observable indicators. Consumption is an indicator given an inversely proportional relationship between consumption and marginal utility, high consumption via low marginal utility or low consumption via high marginal utility. It is also possible to get low consumption and high marginal utility when the investor's other assets perform badly so that low prices are expected for those assets that have consistent covariation with the market portfolio. The capital asset pricing model (CAPM) tells this story[4].

[4]Please refer to Fama and French (1992), (1993) and (1996) for three factor model. This is spoken of in detail in the "asset pricing" text of Cochrane (2001).

We model investors by a utility function defined over current and future values of consumption as $U(C_t, C_{t+1}) = u(C_t) + \delta E_t\left[u(C_{t+1})\right]$. As we know, a gross return is obtained by dividing the payoff by the price. For simplicity, we suppose that the price of the consumption good is unity. An intertemporal decision problem of the two-period optimal consumption for an investor is

$$\underset{C_t, C_{t+1}}{Max}\, U(C_t, C_{t+1}) = u(C_t) + \delta E_t\left[u(C_{t+1})\right]$$

$$s.t. (Y_t - C_t)(1 + r_t) + Y_{t+1} - C_{t+1} = 0$$

where δ is time discount factor, C_t is consumption, Y_t is income and r_t is interest rate between time t and $t+1$. The investor can maximize the intertemporal utility $U(.)$ by the optimal consumption arrangement between the consumption C_t and C_{t+1} subject to the intertemporal budget constraints. At time t, Y_t is known with certainty. All uncertainty is from stochastic Y_{t+1} at time $t+1$ and r_t. The intertemporal budget constraint at time t is the amount of the asset held $A_{t+1} = Y_t - C_t$ while the next period budget constraint, $Y_{t+1} + (1 + r_t)A_{t+1} = C_{t+1}$, is the function of the consumption C_{t+1} at time $t+1$. C_{t+1} changes as Y_{t+1} realizes. In time t, the maximum utility is equal to the sum of the utility, $U(C_t)$, at time t and the expected discounted utility, $U(C_{t+1})$, at time $t+1$. The intertemporal optimal C determines the intertemporal optimal portfolio. See Cochrane (2001) in detail.

Substituting the constrains into the function, and setting the derivative with respect to C_t equal to zero, we obtain the FOC for an optimal consumption and portfolio choice

$$U'(C_t, C_{t+1}) = 0$$

$$\Longleftrightarrow u'(C_t) + \delta E_t u'(\underbrace{(Y_t - C_t)(1 + r_t) + Y_{t+1}}_{\substack{\parallel \\ define\ v(.)}}) = 0$$

$$\Longleftrightarrow u'(C_t) + \delta E_t u'(C_{t+1}) \frac{\partial v}{\partial C_t} = 0$$

$$\Longleftrightarrow 1 = E_t \left[\frac{\delta(1+r_t)u'(C_{t+1})}{u'(C_t)} \right]$$

$$\Longleftrightarrow 1 = E_t \left[(1+r_t) M_{t+1} \right] \tag{3.1}$$

where the variable $M_{t+1} = \frac{\delta u'(C_{t+1})}{u'(C_t)}$ is the stochastic discount factor (or intertemporal marginal ratio of substitution, IMRS). The investor buys more or less of the asset until the FOC holds. Equation (3.1) is the stochastic version of the Euler equation. The Euler equation implies that the investor can maximize utility and optimize by changing consumption and transferring funds between time t and $t+1$, from the lower-utility period to the higher-utility period. Assuming that individual investors can be aggregated into a single representative investor so that aggregate consumption can be used in place of consumption of any particular individual, Equation (3.1) with $M_{t+1} = \frac{\delta u'(C_{t+1})}{u'(C_t)}$, where C_t is aggregate consumption, is known as the consumption CAPM, or CCAPM[5] (Campbell, Lo and Mackinlay, 1997, p304).

We assume that the representative agent in Equation (3.1) maximizes a time-separable power utility function

$$u(C_t) = \frac{C_t^{1-\gamma} - 1}{1-\gamma} \tag{3.2}$$

where γ is the coefficient of relative risk aversion of the investor. As γ approaches one, the power utility function approaches the log utility function

$$u(C_t) = \log(C_t)$$

The first derivative of the power utility function is $u'(C_t) = C_t^{-\gamma}$ for $\gamma \neq 1$ and $u'(C_t) = \frac{1}{C_t}$ for $\gamma = 1$. Substituting the derivative ($\gamma \neq 1$) into Equation (3.1) we get the power-utility-version of the Euler equation

$$1 = E_t \left[(1+r_t)\delta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \right] \tag{3.3}$$

[5]Breeden (1979) develops the CCAPM by defining risk with respect to aggregate consumption.

where the power-utility-version of the stochastic discount factor is

$$M_{t+1} = \delta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma}$$

Using the property $E(AB) = E(A)E(B) + Cov(A, B)$, Equation (3.1) can be written

$$E_t(1 + r_t) = \frac{1 - Cov\{(1 + r_t), M_{t+1}\}}{E_t(M_{t+1})} \tag{3.4}$$

For a risk-free asset, $M_{t+1}$and $(1 + r_{ft})$ are independent because the risk free rate r_{ft} is certain over time so that the covariance between M_{t+1} and the gross return $(1 + r_{ft})$ is zero, $Cov\{(1 + r_t)M_{t+1}\} = 0$. So the risk free rate r_{ft} is

$$1 + r_{ft} = \frac{1}{E_t(M_{t+1})} \tag{3.5}$$

where the gross return $(1 + r_{ft})$ is the reciprocal of the expected stochastic discounted factor. Using the power utility function and setting $R_{ft} = 1 + r_{ft}$, we rewrite Equation (3.5) for $\gamma \neq 1$

$$R_{ft} = \frac{1}{\delta E_t \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \right]} = \frac{1}{\delta E_t \left[\frac{1}{\left(\frac{C_{t+1}}{C_t} \right)^{\gamma}} \right]} = \frac{1}{\delta E_t \left[\frac{1}{\left(\frac{\triangle C}{C_t} + 1 \right)^{\gamma}} \right]}$$

where $u'(C_t) = C_t^{-\gamma}$ and $M_{t+1} = \frac{\delta C_{t+1}^{-\gamma}}{C_t^{-\gamma}}$. R_{ft} is positively associated with consumption growth: rising r_{ft} leads to increasing $\triangle C$ and falling r_{ft} leads to decreasing $\triangle C$. For $\gamma = 1$ Equation (3.5) may be written

$$R_{ft} = \frac{1}{\delta E_t \left[\frac{C_t}{C_{t+1}} \right]} \implies E_t(\frac{C_t}{C_{t+1}}) = \frac{1}{\delta(1 + r_{ft})}$$

where $u'(C_t) = \frac{1}{C_t}$ and $M_{t+1} = \frac{\delta C_t}{C_{t+1}}$. As we know from above that $E_t(\frac{C_{t+1}}{C_t}) = \delta(1+r_{ft})$, we break it into two parts as follows

$$\left(\begin{array}{c} \text{the equation left side: } E_t(\frac{C_{t+1}}{C_t}) = E_t \left[\frac{C_{t+1} - C_t}{C_t} + 1 \right] = E_t(\frac{\triangle C}{C_t} + 1) \\ \text{the equation right side: } \delta(1 + r_{ft}) = \frac{1 + r_{ft}}{1 + \rho} \approx 1 + r_{ft} - \rho \text{ , where } \delta = \frac{1}{1+\rho} \end{array} \right)$$

$$\therefore E_t(\frac{\triangle C}{C_t} + 1) = 1 + r_{ft} - \rho$$

$$E_t(\frac{\triangle C}{C_t}) = r_{ft} - \rho$$

where the relationship between the time discount factor δ and the consumer subjective rate of time preference ρ is defined by $\delta = \frac{1}{1+\rho}$. Consumption change depends on the difference between the risk free rate r_{ft} and the rate of time preference ρ (the Keynes-Ramsey Rule). The consumption growth is positively associated with the risk free rate: higher (lower) risk free interest rates r_{ft} means higher (lower) consumption growth $\triangle C$. Investors will only be "happy" with this situation if interest rates are high. If interest rates were low then investors would try borrowing to increase consumption in today's bad state. It is only high interest rates that stop them from doing this and that makes them "happy" with bad consumption today and good tomorrow. Risk free assets pay a relatively constant gross return R_{ft} and its expected return is the risk free rate. Investors receive relatively low returns when they hold lower-risk assets while investors receive risk premium when they hold riskier assets.

Risk premium is defined by the amount of the money investors receive as compensation for taking riskier assets under more uncertainties. When the investor uses the risky asset, the risk premium for the risk free rate may be written by using Equations (3.4) and (3.5) as follows

$$E_t(1 + r_t) = (1 + r_{ft}) - (1 + r_{ft})Cov\{(1 + r_t), M_{t+1}\}$$

$$\Longleftrightarrow \underbrace{E_t(r_t) - r_{ft}}_{\text{Risk Premium}} = -(1 + r_{ft})Cov\{(1 + r_t), M_{t+1}\} \tag{3.6}$$

The risk premium of the power utility function for $\gamma \neq 1$ is

$$\underbrace{E_t(r_t) - r_{ft}}_{\text{Risk Premium}} = \frac{1}{\underbrace{\delta E_t\left[\underbrace{\frac{1}{\left(\frac{C_{t+1}}{C_t}\right)^{\gamma}}}_{\text{Con. Gro.}}\right]}_{(1) \succ 0}} [-Cov\{(1 + r_t), M_{t+1}\}] \tag{3.7}$$

As shown in Equation (3.7), for the power utility function, risk premium $\{E_t(r_{it}) - r_{ft}\}$ consists of two parts: the consumption growth and the covariance of the return with the stochastic discount factor (or marginal utility). The risk premium is positively associated with the consumption growth and negatively associated with the covariance between $(1 + r_t)$ and M_{t+1}. Generally, $\{E_t(r_{it}) - r_{ft}\} \succ 0$ as long as $-Cov\{(1 + r_t), M_{t+1}\} \succ 0$ which means $Cov\{(1 + r_t), M_{t+1}\} \prec 0$ due to $\delta E_t\left[\frac{1}{\left(\frac{C_{t+1}}{C_t}\right)^{\gamma}}\right] \succ 0$. The higher consumption growth and the more negative the covariance of $(1 + r_t)$ with M_{t+1} is, the higher the risk premium is. Equation (3.6) is the general formula of risk premium in terms of the risky asset return and consumption. The covariance of asset payoffs with consumption drives risk correction to asset prices. In summary, under uncertainty, investors need a higher risk premium to smooth consumption and encourage them to hold riskier assets as they become more risk averse. Otherwise, investors prefer holding the risk free assets and getting lower returns if riskier assets do not offer "attractive" enough risk premium[6].

We rewrite the Euler equation in Equation (3.3) using the property[7] $(1+r_t)\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} =$

[6] Investors could choose insurance and hedges to smooth consumption that happens when risk premium is low while it is not a good hedge when risk premium is high.

[7] $e^{\{\log(1+r_t)+\log\delta-\gamma\log\frac{C_{t+1}}{C_t}\}} = e^{\overbrace{\log\overbrace{(1+r_t)\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}}^{=1}}^{=0}} = e^0 = 1 = (1+r_t)\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}$

$e^{\{\log(1+r_t)+\log\delta-\gamma\log\frac{C_{t+1}}{C_t}\}}$

$$1 = E_t\left[(1+r_t)\delta(\frac{C_{t+1}}{C_t})^{-\gamma}\right] = E_t\left[e^{\{\log(1+r_t)+\log\delta-\gamma\log\frac{C_{t+1}}{C_t}\}}\right]$$

Using the power utility in a lognormal model, we discuss the returns on a risk free asset and a risky asset and give the risk premiums. We assume that the rates of return on assets and consumption growth are jointly lognormal. Using the moment generating function (mgf) of a normal variate Z we can write

$$E(e^Z) = e^{E(Z)+\frac{1}{2}Var(Z)}$$

where $Z = \log(\frac{C_{t+1}}{C_t})^{-\gamma}$ and $e^Z = (\frac{C_{t+1}}{C_t})^{-\gamma}$. Employing the power utility function and taking the log of the Euler equation of risk-free assets in Equation (3.5) with regard to the property $\log E_t[X] = E_t[\log X] + \frac{1}{2}Var_t[\log X]$, we obtain

$$0 = \log\left\{(1+r_{ft})E_t\left[\delta(\frac{C_{t+1}}{C_t})^{-\gamma}\right]\right\}$$

$$0 = \underbrace{\log(1+r_{ft})}_{\text{log return}} + \log\delta + \log\left\{E_t\left((\frac{C_{t+1}}{C_t})^{-\gamma}\right)\right\}$$

$$0 = \underbrace{\log(1+r_{ft})}_{\text{log return}} + \log\delta - \gamma E_t\left[\underbrace{\log(\frac{C_{t+1}}{C_t})}_{\triangle c}\right] + \frac{\gamma^2}{2}Var_t\left[\underbrace{\log(\frac{C_{t+1}}{C_t})}_{\triangle c}\right]$$

$$\underbrace{\log(1+r_{ft})}_{\text{log return}} = -\log\delta + \gamma E_t\left[\underbrace{\log(\frac{C_{t+1}}{C_t})}_{\triangle c}\right] - \frac{\gamma^2}{2}Var_t\left[\underbrace{\log(\frac{C_{t+1}}{C_t})}_{\triangle c}\right]$$

Denoting lowercase letters for logs and setting the log risk free return $\widetilde{r}_{ft} \equiv \log(1+r_{ft})$ and the log consumption growth $c \equiv \triangle c$, we get the risk free interest rate

$$\widetilde{r}_{ft} = -\log\delta + \gamma E_t(\triangle c) - \frac{\gamma^2}{2}\sigma_c^2 \tag{3.8}$$

Equation (3.8) gives the return on a risk free asset. The risk-free return is positively associated with the expected consumption growth and negatively associated with the log of the time discount factor and the volatility of consumption growth. Investors receive a higher return when risk is low, and even more returns as rival investors put off by the risks become more risk-averse. Specifically, when consumption growth is highly volatile, risk-averse investors prefer saving for the next period of consumption, even though their returns are lower. They engage in precautionary saving, saving more and hence expected consumption growth is to offset a more uncertain world in future. Rearranging Equation (3.8), it may be written

$$\begin{aligned} E_t(\triangle c) &= \frac{\log\delta + \log(1+r_{ft})}{\gamma} + \frac{\gamma}{2}\sigma^2_{\triangle c} \\ &= \psi\log(1+r_{ft}) + \psi\log\delta + \frac{\gamma}{2}\sigma^2_{\triangle c} \end{aligned}$$

where ψ is the elasticity of intertemporal substitution, $\psi\gamma = 1$. The expected consumption growth $\triangle c$ is positively associated with the risk free rate of return and its variance. It also implies that investors receive higher consumption growth when they save at a higher rate of return. As investors become more risk-averse, they reduce consumption and save so that future consumption increases. For the return on a risky asset, taking the log of the power utility version of the Euler equation in Equation (3.3), with regard to the property $\log E_t[X] = E_t[\log X] + \frac{1}{2}Var_t[\log X]$, Equation (3.3) becomes

$$0 = E_t\left\{\log\left[(1+r_t)\delta(\frac{C_{t+1}}{C_t})^{-\gamma}\right]\right\} + \frac{1}{2}Var_t\left\{\log\left[(1+r_t)\delta(\frac{C_{t+1}}{C_t})^{-\gamma}\right]\right\}$$

$$\begin{aligned} \iff \quad 0 = E_t&\left[\underbrace{\log(1+r_t)}_{\text{log return}}\right] + \log\delta - \gamma E_t\left[\underbrace{\log(\frac{C_{t+1}}{C_t})}_{\triangle c}\right] \\ &+ \frac{1}{2}Var_t\left[\underbrace{\log(1+r_t)}_{\text{log return}} + \log\delta - \gamma\underbrace{\log(\frac{C_{t+1}}{C_t})}_{\triangle c}\right] \end{aligned}$$

Setting the log return $\widetilde{r}_t \equiv \log(1+r_t)$

$$\Longleftrightarrow 0 = E_t\left(\widetilde{r}_t\right) + \log\delta - \gamma E_t\left(\triangle c\right) + \frac{1}{2}Var_t\left[\widetilde{r}_t + \log\delta - \gamma\left(\triangle c\right)\right]$$

$$\because \text{the } \log\delta \text{ is constant}$$

$$\therefore Var_t\left[\widetilde{r}_t + \log\delta - \gamma\left(\triangle c\right)\right] = Var_t\left[\widetilde{r}_t - \gamma\left(\triangle c\right)\right]$$

$$\Longleftrightarrow 0 = E_t\left(\widetilde{r}_t\right) + \log\delta - \gamma E_t\left(\triangle c\right) + \frac{1}{2}Var_t\left[\widetilde{r}_t - \gamma\left(\triangle c\right)\right]$$

$$\Longleftrightarrow 0 = E_t\left(\widetilde{r}_t\right) + \log\delta - \gamma E_t\left(\triangle c\right) + \frac{1}{2}\left\{Var_t\left(\widetilde{r}_t\right) + \gamma^2 Var_t\left(\triangle c\right) - 2\gamma Cov\left[\widetilde{r}_t, \left(\triangle c\right)\right]\right\}$$

$$\because \text{the log consumption growth } c \equiv \triangle c$$

$$\Longleftrightarrow 0 = E_t\left(\widetilde{r}_t\right) + \log\delta - \gamma E_t\left(\triangle c\right) + \frac{1}{2}\left[\sigma^2_{\widetilde{r}_t} + \gamma^2\sigma^2_c - 2\gamma\sigma_{\widetilde{r}_t,c}\right] \tag{3.9}$$

where the notation $\sigma_{\widetilde{r}_t,c}$ is the unconditional covariance. Equation (3.9) is the log version of the consumption Euler equation. So the return for a risky asset is

$$E_t\left(\widetilde{r}_t\right) = -\log\delta + \gamma E_t\left(\triangle c\right) - \frac{1}{2}\left[\sigma^2_{\widetilde{r}_t} + \gamma^2\sigma^2_c - 2\gamma\sigma_{\widetilde{r}_t,c}\right] \tag{3.10}$$

Subtracting Equation (3.8) from Equation (3.10) yields

$$E_t\left(\widetilde{r}_t\right) - \widetilde{r}_{ft} = -\frac{1}{2}\sigma^2_{\widetilde{r}_t} + \gamma\sigma_{\widetilde{r}_t,c} \tag{3.11}$$

Equation (3.11) gives the risk premium from the log version of the Euler equation. Rearranging Equation (3.11) using the property $\log E_t\left[X\right] = E_t\left[\log X\right] + \frac{1}{2}Var_t\left[\log X\right]$ to deliver $E_t\left(\widetilde{r}_t\right) - \widetilde{r}_{ft}$,

$$\begin{aligned}
E_t\left(\widetilde{r}_t\right) - \widetilde{r}_{ft} &= E_t\left[\log(1+r_t) - \log(1+r_{ft})\right] \\
&= E_t\left[\log\left(\frac{1+r_t}{1+r_{ft}}\right)\right] \\
&= \log E_t\left(\frac{1+r_t}{1+r_{ft}}\right) - \frac{1}{2}Var_t\left[\log\left(\frac{1+r_t}{1+r_{ft}}\right)\right] \\
&= \log E_t\left(\frac{1+r_t}{1+r_{ft}}\right) - \frac{1}{2}Var_t\left[\underbrace{\log\left(1+r_t\right)}_{\widetilde{r}_t} - \underbrace{\log\left(1+r_{ft}\right)}_{\text{constant term}}\right]
\end{aligned}$$

$$= \log E_t\left(\frac{1+r_t}{1+r_{ft}}\right) - \frac{1}{2}Var_t\left(\widetilde{r}_t\right)$$
$$= \underbrace{\log E_t\left(\frac{1+r_t}{1+r_{ft}}\right) - \frac{1}{2}\sigma^2_{\widetilde{r}_t}}_{\text{risk premium}}$$

where $\log(1+r_{ft})$ is constant and $\widetilde{r}_t \equiv \log(1+r_t)$. Putting this rearrangement of $E_t(\widetilde{r}_t) - \widetilde{r}_{ft}$ into Equation (3.11) we rewrite Equation (3.11) in terms of the log version of the expected ratio of the gross returns

$$\log E_t\left(\frac{1+r_t}{1+r_{ft}}\right) = \gamma\sigma_{\widetilde{r}_t,c}$$

As stated above, the risk premium in Equation (3.11) is determined by the coefficient of relative risk aversion multiplied by the covariance of asset returns with consumption growth, where the risk premium is high when the covariance of return with consumption growth is high. For example, an asset that is highly correlated with consumption growth is not very valuable because when times are good (bad) this asset will also become good (bad). This is not purchase for insurance. Investors would prefer to hold an asset that comes good when consumption goes bad and vice versa because it insures investors against hard times. Hence assets positively associated with consumption growth are unattractive and must earn a bigger risk premium. It is noted that the variance of consumption washes out because we subtract the risk free rate and are looking at premia and not to the general level of interest rates. A high variance of consumption will be associated with high levels of all rates in general and is not associated with the risk premium. Risk plays a key role in the expected excess return. When returns positively covary with consumption growth, if consumers are highly risk-averse, they prefer saving not consuming more at the current time in order to consume more in the future. Risk-averse investors probably undertake risky investments when there is high enough risk premium. Otherwise, risk-averse investors will continue to save.

We extend the two-period optimal consumption to the multi-period problem. Following the identical techniques in the two-period optimization, with given available information set at time t and stochastic future income and interest rates, the expected maximum utility of the multi-period optimization problem under uncertainty is

$$\begin{aligned} & \underset{C_{t+i}}{Max}\ U(C_t, C_{t+1}, ..., C_{t+i}) \\ &= Max\ E_t\left[\sum_{i=0}^{\infty}\delta^i u(C_{t+i})\right] \\ &= Max\ \left\{u(C_t) + E_t\left[\sum_{i=1}^{\infty}\delta^i u(C_{t+i})\right]\right\} \\ s.t.\ A_{t+i+1} &= (1+r_{t+i})(A_{t+i} + Y_{t+i} - C_{t+i}), \quad i = 1, 2, 3... \end{aligned}$$

Shifting each term of the intertemporal budget constraint $A_{t+i+1} = (1+r_{t+i})(A_{t+i} + Y_{t+i} - C_{t+i}), \quad i = 1, 2, 3...$ backward i periods $A_{t+1} = (1+r_t)(A_t + Y_t - C_t)$. So the value function at time t is

$$\begin{aligned} V_t(A_t) &= \max_{C_t} E_t\left[\sum_{i=0}^{\infty}\delta^i u(C_{t+i})\right] \\ &= \max_{C_t}\left[u(C_t) + \delta E_t V_{t+1}(A_{t+1})\right] \\ s.t.\ A_{t+1} &= (1+r_t)(A_t + Y_t - C_t) \end{aligned}$$

Substituting the constraint into the objective, and setting the derivative of the utility function $V(.)$ with respect to C_t equal to zero, we obtain

$$V_t'(A_t) = u'(C_t) + \delta E_t\left\{V'_{t+1}\left[\underbrace{(1+r_t)(A_t + Y_t - C_t)}_{(1)\ define\ v(.)}\right]\frac{\partial v}{\partial C_t}\right\} = 0$$

$$u'(C_t) = \delta E_t\left\{V'_{t+1}(A_{t+1})(1+r_t)\right\}$$

Substituting for C_t using the constraint and maximizing with respect to A_t rather than C_t, we get

$$V_t'(A_t) = \frac{\partial V}{\partial A_t} = \delta E_t\left\{V'_{t+1}(A_{t+1})(1+r_t)\ \right\}$$

$$u'(C_t) = V_t'(A_t)$$

Rolling $u'(C_t) = V_t'(A_t)$ forward one period and taking expectations of both sides,

$$E_t\left[u'(C_{t+1})\right] = E_t\left[V_{t+1}'(A_{t+1})\right]$$

Substituting $E_t\left[u'(C_{t+1})\right] = E_t\left[V_{t+1}'(A_{t+1})\right]$ into $u'(C_t) = \delta E_t\left\{V'_{t+1}(A_{t+1})(1+r_t)\right\}$, we obtain

$$u'(C_t) = \delta E_t\left[u'(C_{t+1})(1+r_t)\right]$$

where the marginal cost of saving today equals the expected discounted marginal benefits of consumption tomorrow. The stochastic version of the Euler equation for a multi-period optimization is

$$1 = E_t\left[(1+r_t)M_{t+1}\right] \quad s.t.\ M_{t+1} = \frac{\delta u'(C_{t+1})}{u'(C_t)}$$

In the multi-period problem, all uncertainty is from the stochastic Y_{t+i} and r_{t+i} except Y_t. The intertemporal budget constraint at time $t+i$, $A_{t+1+i} = (1+r_{t+i})(A_{t+i}+Y_{t+i}-C_{t+i})$, means that the assets A_{t+1+i} held in the next period is equal to the returns $(1+r_{t+i})$ multiplied by the reinvestment $(A_{t+i}+Y_{t+i}-C_{t+i})$, which is different from the two-period budget constraint. The difference of $(Y_{t+i}-C_{t+i})$ is reinvested not consumed as it happens in the two-period optimization. Next-period reinvested wealth $A_{t+1+i}+Y_{t+1+i}$, determines next-period consumption C_{t+1+i}. C_{t+i+1} changes as Y_{t+i+1} realizes. At time t, the consumption C_t is constant so that assets A_t held is constant as well if Y_t is given because of $A_t = Y_t - C_t$. When rolling i periods forward A_{t+i} is available if Y_{t+i} is given because C_{t+i} varies as Y_{t+i} realizes. The expected marginal utility of consumption is equal to the marginal value of assets held at time t which is equal to the expected discounted return multiplied by the marginal value of assets held at time $t+1$. It is found that the two-period and multi-period problems have the same Euler equation. The optimization chosen by the two-period (short-term) investor is the same as that of the multi-period (long-term) investor. The multiperiod does not provide more information so that the horizon is not related with the optimization problem. The multi-period investor has become the two-period investor before he enters the two-period horizon because he chooses the optimal consumption and investment and rebalances them at each single period. The investor who is risk averse determines how to make the optimal portfolio decisions. Consequently, optimizing with respect to asset holdings is the same as with respect to consumption so that for the FOC of assets (or asset shares) substituting out consumption is the solution for making optimization.

As mentioned before, an intertemporal optimization is about an optimal consumption and saving decision. The investor makes decisions of investment and consumption simultaneously and smooths consumption over time. In the CCAPM, there is a single representative investor who faces only one aggregated saving and only one non-storable consumption good. The same intertemporal optimization is obtained regardless of whichever (either C_{t+i} or A_{t+i}) the representative investor chooses as the control variable at each period. Saving in each period can help to smooth consumption for the risk-averse investor. Optimal saving achieves the maximal expected returns as optimal consumption does. It is emphasized that, in the CCAPM, the agents do not buy or sell any assets because there is no need to have traded portfolios; nobody else trades with this sole representative investor for the only intertemporal optimization in a one investor-consumer economy.

To better understand the CCAPM, we extend our discussion. We compare the CAPM with the CCAPM.

The CAPM can be derived from the consumption-based model. Cochrane (2001) shows us how to use the log utility of consumption to obtain a CAPM, suggesting that "log utility implies that consumption is proportional to wealth and it's possible to substitute the wealth return for consumption data". We multiply and divide by $VAR(M_{t+1})$ and rewrite Equation (3.4) of the stochastic version of the Euler condition to get a single-beta representation of the CCAPM

$$E_t(1+r_{it}) = \frac{1}{E_t(M_{t+1})} + \left[\underbrace{\frac{COV\left[(1+r_{it}),(M_{t+1})\right]}{VAR(M_{t+1})}}_{\beta_{ic}}\right]\left[\underbrace{-\frac{VAR(M_{t+1})}{E_t(M_{t+1})}}_{\text{risk premium: } E_t(r_{it})-r_{ft}\prec 0}\right]$$

Substituting the risk free rate expressed in Equation (3.5) into the equation showed above, the CCAPM may be written

$$E_t(1+r_{it}) = (1+r_{ft}) + \beta_{ic}\left[E_t(r_{it}) - r_{ft}\right] \tag{3.12}$$

where r_{it} is the rate of return of the risky asset i from the market and β_{ic} is the consumption risk beta. The CCAPM says that with perfect markets the expected return equals the risk free return plus the aggregate consumption beta multiplied by the expected consumption risk premium. The CAPM by Sharpe (1964) and Lintner (1965) is "the first, most famous, and most widely used model in asset pricing". It ties the discount factor $M_{t+1} = \frac{\delta u'(C_{t+1})}{u'(C_t)}$ to the "wealth portfolio". The CAPM is most frequently stated in the expected return-beta version as follows

$$E_t(1+r_{it}) = (1+r_{ft}) + \beta_{im}\left[E_t(r_{mt}) - r_{ft}\right] \tag{3.13}$$

where r_{mt} is the rate of return of the market portfolio m and β_{im} is the market portfolio risk beta. Contrasting Equations (3.12) and (3.13), the point of the CAPM is to avoid the use of consumption data, and use the rate of return on wealth instead. The CAPM links wealth portfolio risk to expected returns. The CAPM is a good measure of risk and thus a good explanation why assets earn higher average returns than others.

Under strong assumptions the CAPM applies period by period. The CAPM exists and prices assets conditional on state variables, which describe the direction of the economy but do not hold unconditionally. As a single-factor model, the CAPM beta does not completely explain the empirical cross section of expected asset returns. Theoretical arguments and empirical evidence suggest that more than one factor is required to characterize the behaviour of expected returns. Multifactor pricing model is considered, naturally. As one of two main theoretical approaches[8], the intertemporal CAPM (ICAPM) developed by Merton (1973) is based on equilibrium arguments. The ICAPM describes "a linear factor model with wealth and state variables that forecast changes in the distribution of future returns or income". The main difference between ICAPM and standard CAPM is additive state variables, which acknowledge the fact that "investors hedge against shortfalls in consumption or against changes in the future investment opportunity set". CAPM and ICAPM ignore consumption decisions. These models assume that investors consume all wealth after one period, or at least that wealth uniquely determines consumption so that preferences defined over consumption are equivalent to preferences defined over wealth. In the real world investors consider many periods when making their investment decisions. In this intertemporal setting the investor must model consumption and portfolio choices simultaneously.

As an intertemporal equilibrium model, the CCAPM is an expansion of the CAPM. The CCAPM factors in consumption as a means of understanding and calculating an expected return on investment. The CCAPM aggregates investors into a single representative agent, who is assumed to derive utility from the aggregate consumption of the economy. The stochastic discount factor in CCAPM is the intertemporal marginal rate of substitution, which is the discounted ratio of marginal utilities in two successive periods for the representative agent. The Euler equations, which are the FOCs for optimal consumption and the portfolio choices of the representative agent, can be used to

[8]Another approach is the arbitrage pricing theory (APT) developed by Ross (1976) is based on arbitrage arguments.

link asset returns and consumption. The CCAPM is helpful for understanding changes of financial asset returns over time, the relationship between saving and consumption and an investor's risk aversion to making an optimal portfolio decision.

We overview some important features of the CCAPM in this section. We give our intuitions: (a) investors expect high returns on almost all assets that are associated with high consumption growth. Investors are "happy" with relatively low consumption today and relatively high consumption tomorrow only when rates are high and vice versa. Put simply, the interest rate is the relative price of consumption today in terms of consumption tomorrow and when this price rises consumption tomorrow is higher too and vice versa. There is also an effect of consumption growth variance on the general level of rates of return, where a more volatile world tends to encourage precautionary saving and investors are happier with lower consumption growth than before for any given rate of return, but this effect washes out of risk premia and just exists in the level of returns; (b) with uncertain rates investors require a risk premia, which is completely different to the effects in (a). Assets that have positive covariance with consumption (growth) are not very useful for diversification and hence must earn higher returns in equilibrium. A bad asset is one whose low consumption tends to be worse due to low returns on the investors' holdings. Assets that have negative covariance with consumption (growth) are very valuable and may even earn a negative risk premium, where bad consumption tomorrow is likely to be offset by high returns because of negative covariance. The covariance of asset payoffs with consumption drives risk correction to asset prices.

3.3 CCAPM with habit persistence

Cochrane's (2001) explanation for the predictability of returns from price/dividend ratios is that "people get less risk averse as consumption and wealth increase in a boom, and more risk averse as consumption and wealth decrease in a recession". Equity premia does not decline as risk aversion increases. There is no way to fix risk aversion to the level of consumption and wealth. The idea is to make "a model in which risk aversion depends on the level of consumption or wealth relative to some trend or the recent past". Following this idea, Campbell and Cochrane (1999) develop the "trend" in consumption using a consumption-based model, while Mcqueen and Vorkink (2004) investigate the "trend" in wealth level in the recent past. We talk about Campbell and Cochrane (1999) in this section and Mcqueen and Vorkink (2004) in the next.

Campbell and Cochrane (1999) emphasize that people's habits for more or less consumption are developed slowly, so that over time the habits form a "trend" in consumption. They specify a habit which is externally determined by the history of aggregate consumption, moves slowly and responds to consumption, and nonlinearly adjusts to the history of consumption. Following Abel (1990), Campbell and Cochrane (1999) employ this external habit for technical convenience[9].

Campbell and Cochrane (1999) start to model an endowment economy with independent and identically distributed (i.i.d.) consumption growth in a lognormal process

$$\triangle c_{t+1} = g + v_{t+1}, \quad v_{t+1} \sim i.i.d.\ N(0, \sigma^2)$$

where log consumption is assumed to follow a random walk with drift g and innovation v_{t+1}. They replace the utility function $u(C)$ with $u(C - H)$ in terms of nonseparable utility over time, where H is the habit level, to maximize the utility function for identical

[9]External habit persistence implies positive serial correlation in consumption changes, which also holds for internal habits. We argue that it does not make much difference to the results for aggregate consumption and asset prices. See Cochrane (2001) for the details.

agents

$$E\sum_{t=0}^{\infty}\delta^t\frac{(C_t-H_t)^{1-\gamma}-1}{1-\gamma}$$

Habits should move slowly in response to consumption and may be written

$$h_t=\phi h_{t-1}+\lambda c_t$$

(Small letters denote the logs of large letters throughout this section e.g. $c_t=\ln C_t$, $h_t=\ln H_t$, etc.)

Campbell and Cochrane (1999) define the surplus consumption ratio $X_t=\frac{C_t-H_t}{C_t}$ to capture the relation between consumption and habit conveniently. They let the surplus consumption ratio of consumption to habit follow an AR(1)

$$x_{t+1}=(1-\phi)\overline{x}+\phi x_t+\lambda\left(x_t\right)\left(c_{t+1}-c_t-g\right)$$

Here, the equation specifies how h responds nonlinearly to c because x is associated with c and h, which means that consumption can never fall below habit since $X=e^x\succ 0$, although it is approximately the same as a traditional habit-formation model[10]. The nonlinear adjustment of habit to consumption guarantees that habit is always below consumption, with finite and positive marginal utility; in other habit models of an endowment economy, habit can be above consumption, with undesirable infinite or negative marginal utility. Campbell and Cochrane (1999) also allow consumption to affect habit differently in different states by featuring a square root type process

$$\begin{aligned}\lambda\left(x_t\right)&=\frac{1}{\overline{X}}\sqrt{1-2\left(x_t-\overline{x}\right)}-1\\ \overline{X}&=\sigma\sqrt{\frac{\gamma}{1-\phi}}\end{aligned}$$

X_t becomes the single state variable in this economy. Time-varying expected returns, price/dividend ratios, etc. are all functions of this state variable.

[10]We call $h_t=\phi h_{t-1}+\lambda c_t$ a traditional habit-formation model. The problem with the traditional model is that it allows consumption to fall below habit, resulting in infinite or imaginary marginal utility.

Campbell and Cochrane (1999) give marginal utility for an external habit as

$$u'(C_t) = u_c(C_t, H_t) = (C_t - H_t)^{-\gamma} = X_t^{-\gamma} C_t^{-\gamma}$$

The external habit, like Abel (1990)'s "catching up with the Joneses" formulation, simplifies analysis and eliminates the terms in marginal utility by which current consumption has an impact on future habits since an individual's habit is determined by "the history of aggregate consumption". With marginal utility, the stochastic discount factor is

$$M_{t+1} \equiv \delta \frac{u_c(C_{t+1}, H_{t+1})}{u_c(C_t, H_t)} = \delta \left(\frac{X_{t+1}}{X_t} \frac{C_{t+1}}{C_t} \right)^{-\gamma}$$

The stochastic process is associated with X and C, and each is lognormal. Campbell and Cochrane (1999) evaluate the risk free rate by evaluating the conditional mean of the stochastic discount factor. The risk free rate is related to the stochastic discount factor by $1 + r_{ft} = 1/E_t(M_{t+1})$ as given by Equation (3.5) in Section 3.2. Taking logs, and using the expressions of x_{t+1} and M_{t+1}, the log risk free rate is

$$r_t^f = -\log E_t(M_{t+1}) = -\log(\delta) + \gamma g - \frac{1}{2}\gamma(1-\phi)$$

Using the basic pricing relation $1 = E_t(M_{t+1}R_{t+1})$ and the definition of returns $R_{t+1} \equiv \frac{P_{t+1}+D_{t+1}}{P_t}$, Campbell and Cochrane (1999) evaluate the price/consumption[11] (or the price/dividend) ratio as a function of the state variable by iteration on a grid

$$\frac{P_t}{C_t}(x_t) = E_t \left[M_{t+1} \frac{C_{t+1}}{C_t} \left(1 + \frac{P_{t+1}}{C_{t+1}}(x_{t+1}) \right) \right]$$

The surplus consumption ratio x_t is the only state variable for the economy, so the price/consumption ratio is a function only of x_t. With the price/consumption ratio, Campbell and Cochrane (1999) can calculate returns, expected returns, the conditional standard deviation of returns, etc.

[11] The price-consumption ratio is an exponentially-weighted average of the expected dividend share. See Cochrane, Longstaff, and Santa-Clara (2003) for the details.

We extend the CCAPM to habit persistence using Campbell and Cochrane (1999). The original motivation of Campbell and Cochrane (1999) was to show that "habit" preferences can generate large equity risk premia but in the following sections we show that these same preferences also lead to ARCH behaviour in asset returns.

3.4 CCAPM with wealth change

Volatility clustering in returns, one of the new important empirical facts in finance, has been involved in most empirical work published in the last decade. There are different explanations for why volatility clustering occurs. These explanations can be divided into two main groups: one proposes an exogenous explanation such as clustered news of economic fundamentals; the other, an endogenous explanation such as heterogeneous traders, where trading process plays the role of triggering volatility autocorrelation. McQueen and Vorkink (2004) argue that clustered economic news explanation of the first group is problematic and the heterogeneous traders explanation of the second group is incomplete. They therefore add a preference explanation, using a preference-based equilibrium asset pricing model to explain low frequency volatility clustering. It is worth noting that there are other explanations using leverage models or state-uncertainty models but both of these inappropriately predict low volatility after good news.

McQueen and Vorkink (2004) mobilize their preference-based equilibrium asset pricing model to explain volatility clustering, drawing upon preference models capable of explaining long-run stock predictability and excess volatility. Their model is structured as an extension of both the preference model of Barberis, Huang and Santos (2001) and the volatility feedback model of Campbell and Hentschel (1992), where the preference model explains features of volatility clustering and the volatility feedback model explains asymmetry in volatility autocorrelation.

According to McQueen and Vorkink (2004) a unique mental scorecard, which records wealth changes and affects investors' level of risk aversion, induces sufficient variation in aversion and sensitivity to news to cause stock volatility. Investors' wealth-varying risk aversion is about loss aversion and scorecard dependence. Investors are more attentive and sensitive to financial news when their expected wealth is perturbed. McQueen and Vorkink (2004) propose a four-stage behavioural process: 1) investors measure their portfolio using the mental scorecard of past investment performance. 2) Investors are more risk averse when their portfolio has an unexpected investment performance. The more risk averse investors are, the more sensitive to news they will be. Stock prices react more to news than when investors are not sensitive. 3) Return shocks cause greater return volatility, which dies out slowly because of investors' persistent attention and sensitivity to news. 4) Investors recover their normal sensitivity to news when they are used to the new level of wealth. Time-varying sensitivity to news endogenously generates clustered returns and volatility clustering and, therefore, state-dependent sensitivity to news is the reason behind volatility clustering in McQueen and Vorkink (2004).

Following Lucas (1978), McQueen and Vorkink (2004) start to maximize expected lifetime utility by allocating wealth between consumption and investment as follows

$$\max \ \mathrm{E}\left[\sum_{t=0}^{\infty}\delta^{t}U\left(C_{t}\right)+b_{0}\overline{C}^{-\gamma}\delta^{t+1}F\left(W_{t+1}\right)\right] \qquad ([1])$$

where thereafter numbers in brackets [ ] denote the numbered equations in McQueen and Vorkink (2004). Financial utility assumptions are made

$$F\left(W_{t+1}\right)=\lambda\left(z_{t},O_{t+1}\right)W_{t+1} \qquad ([3])$$

$$\lambda(z_t, O_{t+1}) = k(a_0 - a_1 z_t) \quad \text{for } O_{t+1} \prec 0 \quad ([4])$$
$$= a_0 - a_1 z_t \quad \text{for } O_{t+1} \succeq 0$$

Notation[12] is given as follows: δ^t - subjective discount rate, C_t - consumption, b_0 - scale parameter of importance of financial utility relative to consumption utility, $\overline{C}$ - aggregate per capita consumption, γ - constant relative risk aversion parameter, W_{t+1} - wealth, z_t - the scorecard, O_{t+1} - return shock, $F(W_{t+1})$ - financial wealth, $\lambda(z_t, O_{t+1})$ - investors' level of risk aversion, k - degree of loss aversion, a_0- investors' baseline level of financial utility derived from gains, a_1 - parameter of how the past performance affecting the magnitude of the utility derived from gains and losses.

In the McQueen and Vorkink (2004) model, investors maximize expected lifetime utility not only from changes in consumption C_t but also from financial wealth $F(W_{t+1})$, where unexpected fluctuations in the value of investors' financial wealth depends on investors' level of risk aversion $\lambda(z_t, O_{t+1})$. The mental scorecard z_t influencing investors' level of risk aversion remembers the prior portfolio shocks as follows:

$$z_{t+1} = \phi z_t + h(z_t) O_{t+1} \quad ([5])$$

where ϕ is a memory parameter $(0 < \phi < 1)$ and $h(z_t)$ is the scorecard's sensitivity to wealth shocks. Return shocks drive changes in the scorecard. When investors' scorecards are perturbed they become more attentive and sensitive to subsequent financial news, which is the unique feature of this model. McQueen and Vorkink (2004) call this the law of motion for z_t.

Taking the first-order conditions of their equation [1] and substituting the marginal utility of financial wealth into the objective[13], McQueen and Vorkink (2004) obtain a

[12] Notation in Section 3.4 is almost identical to that in McQueen and Vorkink (2004) except the one for return shock. In order to make notation homogeneous throughout, we use O_{t+1} to replace X_{t+1} for return shock in McQueen and Vorkink (2004).

[13] Equation [7] in McQueen and Vorkink (2004) is $1 = \delta E_t \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} R_{t+1} + F'(W_{t+1}) \right]$.

wealth-varying risk aversion version of the Euler equation

$$1 = E_t\left[m_{t+1}R_{t+1}\right] \tag{[8]}$$

where the pricing kernel is

$$m_{t+1} = \kappa_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} + \lambda\left(z_t, O_{t+1}\right)\right]$$

$$\kappa_t = \frac{\delta}{1 + \delta E_t\left[\lambda\left(z_t, O_{t+1}\right)\right]E_t\left(R_{t+1}\right)}$$

From their wealth version of the Euler equation, the pricing equation of price-dividend ratios presented is as

$$\frac{P_t}{D_t}\left(z_t\right) = E_t\left[m_{t+1}\left(\frac{P_{t+1}}{D_{t+1}} + 1\right)\frac{D_{t+1}}{D_t}\right] \tag{[10]}$$

McQueen and Vorkink (2004) discuss the model's "qualitative intuition" in which the model has symmetric sensitivity but generates asymmetric responses to news.

McQueen and Vorkink (2004) numerically solve the resulting asset pricing model for the equilibrium, where a solution to the pricing model (Equation [10]) as well as the scorecard's law of motion (Equation [5]) is required. Taking into account the endogenous nature of price-dividend ratios as well as the scorecard, McQueen and Vorkink (2004) use an iterative method and update the pricing model in their equation ([10]) as follows

$$\begin{aligned}\frac{P_t}{D_t}\left(z_t\right) &= E_t\left\{\left(\kappa_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} + \lambda\left(z_t, O_{t+1}\right)\right]\right)\frac{D_{t+1}}{D_t}\left(\frac{P_{t+1}}{D_{t+1}}\left(z_{t+1}\right) + 1\right)\right\} \\ &= \kappa_t E_t\left\{\left[\left(G^{-\gamma}e^{-\gamma\rho\frac{\sigma_v}{\sigma_\epsilon}\varepsilon_{t+1} + \frac{\gamma^2}{2}\left(1-\rho^2\right)\sigma_v^2}\right) + \lambda\left(z_t, O_{t+1}\right)\right]Ge^{\epsilon_{t+1}}\left(1 + \frac{P_{t+1}}{D_{t+1}}\left(z_{t+1}\right)\right)\right\}\end{aligned} \tag{[11]}$$

where $G = ln(g)$. Now the pricing models can be solved numerically by using parameter values given in their Table 1 and conditional moments of return and volatility can be solved by using the price-dividend function and the scorecard's law of motion. McQueen and Vorkink (2004) define the conditional expected stock return as

$$E_t\left(r_{t+1}\right) = E_t\left[\left(\frac{\frac{P_{t+1}}{D_{t+1}}\left(z_{t+1}\right)+1}{\frac{P_t}{D_t}\left(z_t\right)}\right)\frac{D_{t+1}}{D_t}\right] \qquad ([12])$$

which must be solved simultaneously with the price-dividend function and the scorecard law of motion until convergence. Furthermore, based upon the calculations and solutions of conditional expected returns, McQueen and Vorkink (2004) use numerical integration techniques to calculate the expected standard deviation of returns, $E_t\left(\sigma_{t+1}\right)$, in terms of the relevant parameter values given in their Table 1. As shown in their Figures 3 and 4, the expected standard deviation of returns is asymmetrically conditional on the scorecard.

McQueen and Vorkink (2004) conduct simulation practices of monthly returns to investigate if their preference-based equilibrium asset pricing model can explain conditional volatility internally. Following their simulation and sensitivity analysis they argue that their model can generate the consistent conditional volatility found in empirical facts. Moreover, they conclude that their model performs better than either the traditional consumption-only model or the consumption-based model with external habits in Campbell and Cochrane (1999).

McQueen and Vorkink (2004) test the theoretical model of conditional moments. They first test their scorecard's ability to predict conditional volatility. Then they compare their scorecard with other preference scorecards and test its ability to predict conditional excess returns and skewness. Here, we briefly summarize some results relevant to our interests (e.g. conditional volatility and competing scorecards). In the tests of the scorecard's predictive ability of conditional volatility, McQueen and Vorkink (2004) use two regression models with monthly empirical data. The models are run by an estimate of conditional return volatility on an estimate of the lagged scorecard (and on predictions of conditional volatility) as displayed in their Table 3. In the tests of competing scorecards' ability to predict conditional moments, McQueen and Vorkink (2004) compare their scorecard of past investment performance with the scorecard of the log consumption-aggregate wealth ratio in Lettau and Ludvigson (2001) and the scorecard of the surplus consumption ratio in Campbell and Cochrane (1999). They employ another two different empirical regression models, with empirical data at both monthly and quarterly frequencies, as shown in their Table 4. The results (in their Tables 3 and 4) show that their scorecard can predict conditional volatility, excess returns and skewness. It performs better on conditional volatility and skewness than both the scorecard in Lettau and Ludvigson (2001), which is better at predicting excess returns, and the scorecard in Campbell and Cochrane (1999). McQueen and Vorkink (2004) conclude that the preference-based equilibrium asset pricing model, in which the utility is obtained from consumption and wealth changes, is capable of explaining many stylized facts pertaining to conditional volatility, even the new empirical facts in finance including excess returns, high risk premium and skewness.

We extend the CCAPM to derive utility from wealth changes using McQueen and Vorkink (2004), which is able to explain the conditional volatility found in US stock data. McQueen and Vorkink (2004) employ a preference-based asset pricing model to capture long-term stock predictability and excess volatility. The model includes wealth-varying degrees of risk aversion and sensitivity to news. They show that the mental scorecard recording the market's sensitivity to news and affecting the agents' level of risk aversion, due to wealth changes and experience based loss aversion, is able to explain conditional volatility - even its asymmetric properties. The original motivation of McQueen and Vorkink (2004) was to stress the fact that revisions to wealth introduced in the utility function can lead to ARCH behaviour.

3.5 Model

3.5.1 Background

The aim of the chapter is to investigate whether the Lucas two-country monetary model with habit in Moore and Roche (2006) can generate generic predictable conditional volatility in spot returns $\triangle s_{t+1}$ not in the risk premium, or foreign asset returns, or in asset returns in general.

In CCAPM, under one non-storable consumption good, a single representative consumer's aggregate consumption becomes equal to the total economy consumption, so that total expected consumption (growth) in the economy is linked to expected returns. If there is no capital stock and there are no perishable consumption goods, as in the Lucas model, then things become even simpler, with individual consumption becoming equal to economy wide consumption equal to economy wide exogenous income.

Before we show how the Lucas two-country monetary model with habit may generate predictable ARCH, we discuss the ability of the single Lucas model to produce conditional volatility.

Lucas (1982) sets up a dynamic general equilibrium model of an endowment economy with a complete market. The Lucas model prices foreign exchange that depends on preference. In the Lucas model, representative agents in two countries are provided with identical preferences for two consumption goods but with different stochastic endowments of these. The assumption of the Lucas model is that securities markets are complete so that there is complete pooling of risks. With identical preferences, agents will consume exactly one half of the endowment of each good in each period and maximize their expected infinite utility function in each country.

We discuss our initial ideas. Since the Lucas model in its simplest form can never generate ARCH effects in spot returns s_t, the real exchange rate in any Lucas model is just the relative price of home to foreign goods. In competitive models, this relative price is always equal to the ratio of the goods' marginal utilities (the marginal utility of one good divided by the marginal utility of the other). Hence we find that for domestic good L and foreign good F

$$q_t \equiv \frac{S_t P_t^*}{P_t} = \frac{\partial U(.)/\partial L_t}{\partial U(.)/\partial F_t}$$

where t is time; q_t is the relative price of home to foreign goods; S_t is spot rate; P_t is the price of domestic good; P_t^* is the price of foreign good; ∂ is the derivative and $U(.)$ denotes utility. Note that we have not made any assumptions about the time separability of U yet. So this condition is true for all types of the Lucas model, whatever the form of $U(.)$ is. We denote the lower case as logs and the upper as levels. Now we can use the cash in advance constraints, which dictate that all $L(F)$ goods must be paid for with domestic(foreign) money $M(N)$, where $M_t = P_t L_t$ and $N_t = P_t^* F_t$, to substitute out for prices. In terms of monies, we get the condition

$$q_t \equiv \frac{S_t N_t L_t}{M_t F_t} = \frac{\partial U(.)/\partial L_t}{\partial U(.)/\partial F_t}$$

We now solve for S in terms of all the other conditions

$$S_t = \frac{M_t F_t}{N_t L_t} \frac{\partial U(.)/\partial L_t}{\partial U(.)/\partial F_t}$$

This equation shows that the exchange rate depends on the relative monies but also depends on the marginal utilities and the home and domestic (exogenous) endowment streams arising from the Lucas "trees".

There are several potential ways to get ARCH in S. One is to fix the relative money processes to make them have ARCH. But by doing this we would be putting ARCH in to get ARCH out. It is not reasonable. Another way is to fix the U or u functions in such a way as to make their derivatives have ARCH time series properties.

First, if the simple Lucas model is taken to be Lucas plus standard time separable power utility, then the marginal utility ratio is just a simple function of the ratio of home to foreign consumptions C_{l_t}/C_{f_t} and this ratio is just l_t/f_t because of the endowment economy, under which all outputs must be consumed. Substituting this into the formula for S it is found that S is just proportional to relative money supplies, exactly as in the simple "ad hoc" monetary model. The simple Lucas model cannot get ARCH for S out of this unless we assume ARCH for monies, which we already know is counterfactual.

However, if we specify an exotic U, for example, which depends on a "habit", a ratio of marginal utilities is obtained to give an ARCH behaviour. The habit defined by Campbell and Cochrane (1999) to solve the equity premium puzzle in Section 3.3 is an AR(1). The lagged level of consumption is the "shock" in the AR(1) process. Solving the AR(1) backwards gives the habit H as something like

$$h_t \approx \frac{1-\phi}{\phi}\sum_{i=1}^{\infty}\phi^i c_{lt-i}$$

which is equivalent to $h_t = \phi h_{t-1} + \lambda c_t$ given in Section 3.3[14]. This seems to be a reasonable assumption. Now instead of U being in terms of actual consumption it is written in terms of consumption relative to habit - i.e. the surplus consumption ratio of the form $\frac{C_{lt}-H_t}{C_{lt}}$. It turns out that if the habit H is as above then surplus consumption will be an AR(1) but with a sensitivity to shocks parameter λ whose value depends on the previous level of surplus consumption. This is the Moore and Roche (2006) surplus consumption evolution equation.

[14]See Campbell, Lo, and Mackinlay (1997), p330-331, for the details.

3.5.2 A Lucas two-country monetary model with habit

Moore and Roche (2006) employ the Lucas (1982) two-country monetary model with an external habit persistence devised by Campbell and Cochrane (1999) to solve exchange rate puzzles (disconnect, forward bias, and Meese-Rogoff) and mimic the unconditional volatilities of real and nominal exchange rates, forward premium, expected spot returns and expected forward profits.

Moore and Roche (2006) assume that consumption growth and money growth follow an AR(1) processes

$$\Delta c_{t+1}^{j} = (1-\rho_{\mu})\overline{\mu} + \rho_{\mu}\Delta c_{t}^{j} + v_{t+1}^{j}\ , \quad v_{t+1}^{j} \sim N(0,\sigma_{v}^{2})\ , \quad j=1,2 \qquad ([17],3.14)$$

$$\Delta m_{t+1}^{j} = (1-\rho_{\pi})\overline{\pi} + \rho_{\pi}\Delta m_{t}^{j} + u_{t+1}^{j}\ , \quad u_{t+1}^{j} \sim N(0,\sigma_{u}^{2})\ , \quad j=1,2 \qquad ([18],3.15)$$

where $\overline{\mu}$ ($\overline{\pi}$) is the unconditional mean of consumption (money) growth; v_{t+1}^{j} (u_{t+1}^{j}) are the shocks to consumption (money) growth while the shocks to consumption and money growth are uncorrelated; σ_v^2 (σ_u^2) is the variances of shocks to consumption (money) growth. We write, next to our numbered equation in ( ), a bracket [ ] in which the number written corresponds to the serial number of the equation in Moore and Roche (2006). All equations with identical notation in this section are citations from Moore and Roche (2006).

Moore and Roche (2006) define habit persistence using an aggregate consumption externality. They give the maximized utility function (assuming identical parameters for both countries) as

$$\sum_{t=0}^{\infty}\beta^{t}\left\{\frac{\left(C_{it}^{1}-H_{it}^{1}\right)^{1-\gamma}}{1-\gamma}+\frac{\left(C_{it}^{2}-H_{it}^{2}\right)^{1-\gamma}}{1-\gamma}\right\}, \quad i=1,2 \qquad ([5],3.16)$$

$$s.t.\ W_{t+1} = S_{t+1}B_{t}^{2} + B_{t}^{1} + \ P_{t}^{1}Y_{t}^{1} \qquad ([6])$$

and the wealth constraint as

$$W_t = P_t^1 C_t^1 + S_t P_t^2 C_t^2 + q_t^1 B_t^1 + S_t q_t^2 B_t^2 \tag{[7]}$$

where β is the discount factor, γ is a curvature parameter, C_{it}^j is the consumption of goods and services of country j by the household of country i, H_{it}^j is the subsistence consumption / habit of goods and services of country j by the household of country i, W_{t+1} is the next-period wealth, S_t is the level of the spot exchange rate, B_t^j is the amount of one-period discount bonds from country j, P_t^j is the price level in country j, Y_t^j is the endowment in country j and q_t^j is the nominal bond price in country j. As shown above, the next-period wealth consists of a monetary transfer $(S_{t+1}B_t^2)$, dividends (B_t^1) and market value of securities $(P_t^1 Y_t^1)$ while the three parts in the wealth constraint are goods $(P_t^1 C_t^1 + S_t P_t^2 C_t^2)$, equity $(q_t^1 B_t^1)$ and a money transfer $(S_t q_t^2 B_t^2)$.

Moore and Roche (2006) assume that the cash-in-advance constraint is

$$\frac{M_t^j}{P_t^j} = C_t^j \ , \qquad j = 1, 2 \tag{[8]}$$

and define the surplus consumption ratio (SCR) as

$$X_t^j = \frac{\overline{C}_t^j - H_t^j}{\overline{C}_t^j} \ , \qquad j = 1, 2 \tag{[9]}$$

where X_t^j is the SCR of country j, M_t^j is the money in country j and $\overline{C}_t^j$ is the aggregate consumption per capita of goods and services in country j.

Moore and Roche (2006) let the log of the surplus consumption ratios follow an AR(1) process

$$x_{t+1}^j = (1-\phi)\overline{x} + \phi x_t^j + \lambda(x_t^j)(v_{t+1}^j) \ , \quad j = 1, 2 \tag{[10],3.17}$$

where $\phi(< 1)$ is the habit persistence parameter, $\overline{x}$ is the steady state value for the logarithm of the surplus consumption ratio.

Moore and Roche (2006) also allow consumption to affect habit differently throughout states by featuring a squared root type process as in Campbell and Cochrane (1999), where they define the sensitivity function $\lambda(x_t^j)$ of the log surplus consumption ratio to endowment innovations to non-linearly depend on the current log surplus consumption ratio.

$$\lambda(x_t^j) = \frac{\sqrt{1-2\left(x_t^j - \overline{x}\right)}}{\overline{X}} - 1 \quad for\ x_t^j \leq x_{\max} \qquad j = 1, 2$$

$$= 0 \qquad for\ x_t^j \succ x_{\max} \qquad ([11],3.18)$$

$$where\ \ x_{\max} = \overline{x} + \frac{1-\overline{X}^2}{2} \qquad \text{and} \qquad \overline{X} = \frac{\gamma\sigma_v}{\sqrt{\gamma(1-\phi)-\delta}} \qquad ([12])$$

where $\overline{X}$ is the steady state value of the surplus consumption ratio, δ is the parameter in steady state surplus consumption and $\gamma(1-\phi)-\delta > 0$.

Using the FOCs the optimization problem is solved and Moore and Roche express the nominal exchange rate

$$S_t = \frac{\left(C_t^2\right)^{1-\gamma}\left(X_t^2\right)^{-\gamma}}{\left(C_t^1\right)^{1-\gamma}\left(X_t^1\right)^{-\gamma}} \frac{M_t^1}{M_t^2} \qquad ([A10],3.19)$$

The log of the nominal exchange rate is

$$s_t = -(1-\gamma)(c_{t+1}^1 - c_{t+1}^2) + \gamma(x_{t+1}^1 - x_{t+1}^2) + (m_{t+1}^1 - m_{t+1}^2) \qquad ([20],3.20)$$

The volatility of nominal exchange rates is given by the variance of spot returns

$$Var(s_{t+1} - s_t) = 2\left[\frac{\sigma_u^2}{1-\rho_\pi^2} + (1-\phi)^2\gamma^2\sigma_x^2 + \sigma_v^2\left(1 - \frac{\gamma}{\overline{X}}\right)^2\right] \qquad ([21],3.21)$$

where σ_x^2 is the variance of the surplus consumption ratio. Equation ([21],3.21) is helpful when using the moment expressions to provide some insight into the simulated results in Moore and Roche (2006).

3.5.3 Implications of the model

We aim to investigate the model's ability to generate conditional volatility in spot returns. Hence, we explore the properties of spot returns Δs_t and calculate innovations $\widehat{\zeta}_t$. We analyze the theoretical model (Moore and Roche's model) and then derive an implied ARMA(2,2) process of spot returns Δs_t. We filter the data for removing AR components to obtain the filtered spot returns Δs_t^f. We calculate innovations $\widehat{\zeta_t}$, subject to conditional volatility, after estimating an MA(2) model for the filtered spot returns Δs_t^f.

Using Moore and Roche's equation (10), of an AR(1) log surplus consumption ratio, and equation (A12), of the change in spot rates s, we analyze the ARMA(2,2) process of spot returns.

We calculate the difference between country one and two in surplus consumption ratios using Equation ([10],3.17)

$$x_t^1 - x_t^2 = \phi(x_{t-1}^1 - x_{t-1}^2) + \lambda_{t-1}^1 v_t^1 - \lambda_{t-1}^2 v_t^2 \tag{3.22}$$

We also take the difference between country one and two in money growths to be represented by Equation ([18],3.15) where both money growths are about an AR(1) with the identical AR(1) coefficient ρ_π.

$$\Delta m_t^1 - \Delta m_t^2 = \rho_\pi(\Delta m_{t-1}^1 - \Delta m_{t-1}^2) + u_t \tag{3.23}$$

Denote

$$x_t = x_t^1 - x_t^2$$

$$m_t = \Delta m_t^1 - \Delta m_t^2$$

$$w_t = \lambda_{t-1}^1 v_t^1 - \lambda_{t-1}^2 v_t^2$$

where $var(w_t) = \sigma^2_{wt}$ is conditionally heteroscedastic but $E(w_t w_{t-i}) = 0$ for $i = 1, 2, 3.....$ and u_t is *iid*. Now Equations (3.22) and (3.23) may be written as

$$x_t = \phi x_{t-1} + w_t \tag{3.24}$$

$$m_t = \rho_\pi m_{t-1} + u_t \tag{3.25}$$

Using the lag operator and noting that the roots are less than one, therefore invertible, we rewrite Equations (3.24) and (3.25) respectively as

$$x_t = \frac{w_t}{1 - \phi L}$$

$$m_t = \frac{u_t}{1 - \rho_\pi L}$$

The process of spot returns described in Moore and Roche's equation (A12) is

$$\begin{aligned} \Delta s_{t+1} &= \left(\Delta m^1_{t+1} - \Delta m^2_{t+1}\right) - \gamma(1-\phi)\left(x^1_t - x^2_t\right) \\ &\quad - \left\{1 - \gamma\left[1 + \lambda\left(x^1_t\right)\right]\right\} v^1_{t+1} + \left\{1 - \gamma\left[1 + \lambda\left(x^2_t\right)\right]\right\} v^2_{t+1} \end{aligned} \tag{[A12]}$$

Here, we write down the form for time t instead of Moore and Roche's time $t + 1$ and thus, in this notation, Moore and Roche's equation (A12) becomes

$$\Delta s_t = m_t + a x_{t-1} + z_t \tag{3.26}$$

where $var(z_t) = \sigma^2_{zt}$ is related to w_t above and it is likewise conditionally heteroscedastic but with $E(z_t z_{t-i}) = 0$ for $i = 1, 2, 3...$, and $a = -\gamma(1 - \phi)$.

Substituting x in Equation (3.24) and m in Equation (3.25) into Equation (3.26), we get

$$\Delta s_t = \frac{u_t}{1 - \rho_\pi L} + a\frac{w_{t-1}}{1 - \phi L} + z_t \tag{3.27}$$

Multiplying both sides by $(1 - \rho_\pi L)(1 - \phi L)$ gives

$$(1 - \rho_\pi L)(1 - \phi L)\Delta s_t = (1 - \phi L)u_t + a(1 - \rho_\pi L)w_{t-1} + (1 - \phi L)(1 - \rho_\pi L)z_t$$

It is noted that the error structure on the right hand side (RHS) is an MA(2) and because we have an AR(2) structure on the left hand side (LHS) the s series is an ARMA(2,2). The composite error has no autocovariances above 2 so it can be therefore written as

$$(1 - \phi L)u_t + a(1 - \rho_\pi L)w_{t-1} + (1 - \phi L)(1 - \rho_\pi L)z_t = \zeta_t + \theta_1\zeta_{t-1} + \theta_2\zeta_{t-2}$$

where ζ_t is (highly) conditionally heteroscedastic and $E(\zeta_t\zeta_{t-i}) = 0$ for all $i = 1, 2, 3, ...$

Using Moore and Roche's calibrated baseline parameters, the values of $\rho_\pi = 0.1$ and $\phi = 0.999$, and our simulated series for s, we can compute $(1 - \rho_\pi L)(1 - \phi L)\Delta s_t$ directly. We call this series Δs_t^f, where "^{f}" denotes filtered. Then we have

$$\Delta s_t^f = \zeta_t + \theta_1\zeta_{t-1} + \theta_2\zeta_{t-2} \tag{3.28}$$

This series is an MA(2) in a conditional heteroscedastic error ζ. We can use the Method of Moments to get good estimates of θ_1 and θ_2. Then denoting estimates by $\widehat{}$ we can compute the genuine innovation in the change in s as

$$\widehat{\zeta_t} = (1 + \widehat{\theta}_1 L + \widehat{\theta}_2 L^2)^{-1}\Delta s_t^f \tag{3.29}$$

As shown above, the change in spot rates, Δs_t, is an ARMA(2,2) and the filtered series for s, Δs_t^f, is an MA(2) in the conditional heteroscedastic error ζ, which we call the properties implied by the theoretical model[15] hereafter. We calculate innovation $\widehat{\zeta}_t$ with the details given in Appendix B.1. We now have a series of innovation $\widehat{\zeta}_t$ with which we can apply the estimations of the ARCH class models' to directly. We estimate the best fitting ARCH class model(s) to innovation $\widehat{\zeta}_t$. The model is able to explain conditional volatility in spot returns as long as innovation $\widehat{\zeta}_t$ has ARCH performance. It is emphasized that the innovation $\widehat{\zeta}_t$ in the ARMA(2,2) process for spot returns is by definition conditionally heteroscedastic. The data we will simulate for spot returns and for the innovation in the spot returns process will definitely be conditionally heteroscedastic. We give the proof in the next section.

3.6 Solution and evaluation

We now turn our focus to presenting the baseline simulation results. First, we check if the simulations[16] in the baseline for the level of exchange rates (S) have the same time series properties as in Moore and Roche (2006). Second, we evaluate the efficiency of the simulations intended to capture the implied properties of the theoretical model. We also assess sensitivity to parameter changes.

We numerically solve the model. We employ the quarterly calibration parameter values assumed in the baseline framework in Moore and Roche (2006), where the parameters β and δ are chosen from the literature; γ and ϕ are chosen to make sure that the surplus consumption ratio ($\overline{X}$) is approximately 5% and that the value of local risk aversion is no more than 10; the parameters of endowment and money growth are chosen by using US data. Beside the parameterization in the baseline framework, in the sensitivity analysis, we set $\gamma = 0.7$, $\delta = -0.0025$, $\rho_\pi = 0$ or $\phi = 0.995$, respectively.

[15] Moore and Roche's model in Moore and Roche (2006).

[16] We code in Matlab.

The baseline's parameterization is displayed in Table 3.1.

Parameter	**Variable**	**Value**
Discount factor	β	0.99
Curvature of the utility function	γ	0.5
Parameter in steady state surplus consumption	δ	-0.005
AR(1) coefficient of log surplus consumption	ϕ	0.999
AR(1) coefficient of money growth	ρ_π	0.1
AR(1) coefficient of consumption growth	ρ_μ	0.00
Unconditional mean of money growth at steady state	$\overline{\pi}$	0.0136
Unconditional mean of consumption growth at steady state	$\overline{\mu}$	0.004725
Standard deviation of money growth	σ_u	0.00946
Standard deviation of consumption growth	σ_v	0.0075
Steady state value of surplus consumption ratio	$\overline{X}$	0.0506
Log steady state value of surplus consumption ratio	$\log(\overline{X})$	-2.9845
Max value of surplus consumption ratio	$X_{\max}$	0.0833
Max value of log surplus consumption ratio	$x_{\max}$	-2.4858
Local Relative Risk Aversion ($\preceq 10$)		9.88826

Notes: All parameters are quoted from Moore and Roche (2006).

Table 3.1: Model parameter values in baseline

Correct simulation techniques are guaranteed to construct accurate exchange rates and spot returns. The simulation procedure is executed as follows:

1. to generate a time series of consumption growth ($\triangle c$) using Equation ([17],3.14), where we set initial consumption growth at its steady state value as given in Table 3.1;

2. to generate a time series of money growth ($\triangle m$) using Equation ([18],3.15), where we set initial money growth at its steady state value as given in Table 3.1;

3. to accumulate the variables generated in step 1 and 2 so as to obtain the log level;

4. to generate a time series of log surplus consumption ratio ($x = \log X$) using Equations ([10],3.17) and ([11],3.18), where we set initial log surplus consumption ratio x at $\overline{x}$ and initial sensitivity function $\lambda(x)$ at $\lambda(\overline{x})$;

5. to construct a time series of exchange rates S using Equation ([A10],3.19);

6. to repeat steps 1-5 for each series that is constructed in the baseline framework and sensitivity analysis, respectively.

Before we investigate, examine and evaluate the theoretical model's implied properties, we carry out two simulation exercises. We start to simulate several time series (132 data points) with approximately the same properties of volatility and persistence as the exchange rates in the first-differenced data and the spot returns found in Moore and Roche's Tables 9 and 10. We compare the statistics in the theoretical economy to those of the empirical data. The aim of the first exercise is to assess our simulation techniques by (more or less) replicating the results in Moore and Roche (2006). We also simulate several longer series (10000 data points) to capture the Moore and Roche model's properties. The motivation behind the latter is precision: Moore and Roche simulate only 132 observations and so the estimates of their model's properties are very imprecise. We are willing to deal with imprecision when we estimate the properties of the real world data, where only 132 observations are available, as in Moore and Roche (2006). However, we can generate longer series to reduce imprecision when estimating the model's properties. Hence the parameter values, variances, etc. estimated in the second exercise will be closer to the model's true parameters, variances, etc.

3.6.1 Simulation

Simulation 1

We simulated the model 1000 times generating log surplus consumption ratio ($x = \log X$), money growth ($\triangle m$) and consumption growth ($\triangle c$) for 132 observations[17]. After producing x, m and c we can construct exchange rates S.

[17] We generate log surplus consumption ratio (x), money growth ($\triangle m$) and consumption growth ($\triangle c$) for 232 observations and discard the first 100 observations for each series in the baseline and sensitivity analysis, respectively.

We report the statistics of volatility and persistence for both the spot returns and the first-differenced exchange rates in Table 3.2. We find the approximately consistent results by comparing Moore and Roche's results to ours. The statistics of the first-differenced (FD for short afterwards) data are reported in Panel A of Table 3.2, where we filter the logarithm of the simulated exchange rates, $log(S)$, using the FD filter. First, for the property of volatility, in the baseline, the mean of volatility of the FD $log(S)$ in our case is 5.75, which is close to Moore and Roche's 6.36 and is much closer to the empirical value, 5.09, while the std. dev. of volatility in our case, 0.080, is less than Moore and Roche's 0.232. We also look at one other case apart from the baseline. In $\delta = -0.0025$ of the sensitivity analysis, the mean of volatility in our case is 5.51, which is similar to 5.38 in Moore and Roche (2006) and is much closer to 5.09 of the empirical data and even is better than our baseline's 5.75. At the same time, in $\delta = -0.0025$, the std. dev. of volatility in our case is 0.054, which is less than 0.143 in Moore and Roche (2006). We find that, in the sensitivity analysis, volatility increases if we raise the parameter value of γ and the absolute value of δ, respectively, or decrease the parameter values of ϕ and ρ_π, respectively, which is consistent with what is found in Moore and Roche (2006)[18].

Second, we find that the persistence of our simulated data is approximately consistent with that of both the empirical and the theoretical data in Moore and Roche (2006). Our simulated data is a slightly negatively autocorrelated while the autocorrelation parameters of the theoretical data in Moore and Roche (2006) are positive or negative. However, both the absolute values tend to be approximately around the value point at 0.02. We also find, but do not report, the consistent high persistence of the exchange rates filtered by using the Hodrick-Prescott filter, as in Moore and Roche

[18] Moore and Roche (2006) report the same mean of volatility in the baseline and $\rho_\pi = 0$, which is still consistent with our results.

Panel A
Properties of exchange rates in first-differenced data

	Empirical Data	Simulated Data									
		Moore and Roche (2006)					Liu (2007)				
		Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
Volatility (%)											
Mean	5.09	6.36	7.42	5.38	6.36	8.48	5.75	7.59	5.51	6.52	8.86
Std.Dev	1.705	0.232	0.203	0.143	0.232	0.233	0.080	0.070	0.054	0.087	0.081
Persistence											
Mean	0.07	0.02	0.01	0.02	-0.02	-0.01	-0.04	-0.02	-0.02	-0.05	-0.03
Std. Dev	0.051	0.004	0.004	0.004	0.004	0.004	0.158	0.135	0.134	0.156	0.130

Panel B
Properties of spot returns

	Empirical Data	Simulated Data									
		Moore and Roche (2006)					Liu (2007)				
		Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
Volatility (%)											
Mean	5.09	6.36	7.42	5.38	6.36	8.48	5.75	7.59	5.51	6.52	8.86
Std. Dev	1.705	0.232	0.203	0.143	0.232	0.233	0.080	0.070	0.054	0.087	0.081
Persistence											
Mean	0.07	0.02	0.01	0.02	-0.02	-0.01	-0.04	-0.02	-0.02	-0.05	-0.03
Std. Dev	0.051	0.004	0.004	0.004	0.004	0.004	0.158	0.135	0.134	0.156	0.130

Notes: Empirical data refers to nine real exchange rates (CAD/USD, GBP/USD, JPY/USD, CHF/USD, EUR/USD, CAD/EUR, GBP/EUR, JPY/EUR, and CHF/EUR) obtained from DataStream International in Moore and Roche (2006). See Moore and Roche's table 2 for notes to empirical data. Statistics of empirical data are calculated basing on the displayed numbers in Moore and Roche (2006)'s table 3 and 4. The mean of empirical data is the average of nine exchange rates. The Std. Dev of empirical data is the statistic of the series of nine exchange rates. The statistics of simulated data of Moore and Roche (2006) are quoted directly from Moore and Roche (2006)'s table 9 and 10. The statistics of our simulated data are given under the title of "Liu (2007)". The standard deviation (std dev for short) reported in theoretical economy are the averages from the 1000 simulations. Volatility is measured by the std dev and persistence is measured by the first-order autocorrelation coefficient.

Table 3.2: Statistics of empirical data vs. theoretical data

(2006). Our findings suggest a later application of a GARCH class model to conditional volatility. In GARCH specifications, the autoregressive root which governs the persistence of volatility shocks is the sum of the ARCH parameter[19] (α) plus the GARCH parameter (β). When this root is very close to unity volatility, shocks are quite persistent and die out rather slowly.

We report the properties of spot returns in Panel B of Table 3.2. We find the consistent results, which are compared to those in Moore and Roche (2006). We also find that the statistics of spot returns in Panel B are the same as those of the first-differenced exchange rates in Panel A due to the same log first-order difference process.

In general, we (approximately) replicate the results of Moore and Roche (2006). The data we simulated meet Moore and Roche's statistical criteria[20] and have (approximately) the same time series properties as theirs. Moore and Roche use the moment expressions in Equation ([21],3.21)[21] to provide some insight into their simulated results. The volatility of the fundamentals is able to explain the volatility in the nominal exchange rates.

Simulation 2

We simulated the model once for 10000 observations to investigate the model's implied properties: Δs_t is an ARMA(2,2); Δs_t^f is an MA(2); Δs_t and its innovation

[19]See notes in Chapter 2 for notations of GARCH class conditional volatility models.

[20]We thank Prof. Michael Moore and Dr. Maurice Roche for their kind help. We check our simulation by using their statistics. We report the relevant statistics of our simulation: 1) the std. dev. of log X_{home} (log $X_{foreign}$) is 30.27% (27.24%) while Moore and Roche's std. dev. of $logX$ is about 25%; 2) the std. dev. of the change in log X_{home} (log $X_{foreign}$) is 7.63% (6.82%) while Moore and Roche's std. dev. of the change in $logX$ is be about 6.5%; 3) the std. dev. of (home and foreign) consumption growth is 0.75% while Moore and Roche's std. dev. of consumption growth is 0.75%. It is exogenous. Moore and Roche suggest that the std. dev. of the change in log surplus consumption is at least 10 time that of consumption growth; 4) the log of the surplus consumption ratio is always negative and the level of the surplus consumption ratio is always non-negative, which is consistent with Moore and Roche (2006) and Campbell and Cochrane (1999). The reason is "in the continuous-time limit, the x_t process never attains the region $x > x_{\max}$". So the log surplus consumption ratio is always negative since $x \prec x_{\max} = -2.4858$ as in Table 3.1.

[21]The expressions in Equation ([21],3.21) are approximations based on Moore and Roche's equation (A14).

ζ are conditional heteroscedastic. The time series of spot returns is constructed by using the simulated 10000 quarterly exchange rates S. We investigate the time series of spot returns to the model's implied properties. Innovations of spot returns are subject to conditional volatility. In the baseline, all initial values and parameterization are the same as in Simulation 1 using Table 3.1. We repeat this exercise for each of the sensitivity variants to see the implications that this has on the conditional volatility of the model. The testing results[22] are reported in Tables 3.3-3.5.

$$\Delta s_t^f = c_0 + \theta_1 \zeta_{t-1} + \theta_2 \zeta_{t-2}$$

		Sensitivity analysis			
	Baseline	$\gamma = 0.7$	$\delta = -0.0025$	$\rho(\pi) = 0$	$\phi = 0.995$
Panel A: parameter estimates					
c_0	-8.44E-07	-1.61E-06	8.53E-07	4.76E-07	1.24E-06
	(-0.319)	(-0.443)	(0.395)	(0.194)	(0.347)
θ_1	-0.7852*	-0.8088*	-1.1354*	-0.7173*	-0.7851*
	(-84.638)	(-85.936)	(-114.597)	(-82.015)	(-84.469)
θ_2	-0.2112*	-0.1878*	0.1376*	-0.2792*	-0.2114*
	(-22.775)	(-19.940)	(13.884)	(-31.926)	(-22.743)
Panel B: adjusted R-squared ($\overline{R}^2$)					
Lag 1-2	0.50	0.51	0.57	0.45	0.50
Lag 3-10	0.011	0.002	0.006	0.011	0.012

Notes: In Panel A, t-ratios in parentheses; * denotes significance at the 1% level; in Panel B, adjusted R-squared is generated by regressing the filtered spot returns on the lags 1-2 and on the lags 3-10 of the conditional heteroscedastic errors, respectively.

Table 3.3: The filtered spot returns of an MA(2) process

Table 3.3 shows that Δs_t^f is an MA(2) in both the baseline and sensitivity analysis. The estimates of the constant, θ_1, and θ_2 in Equation (3.28) are displayed in Panel A of Table 3.3. All the coefficients of θ_1 and θ_2 are highly significant (at 1%) with the non-significant constants in both the baseline and sensitivity analysis. Furthermore, in order to prove that Δs_t^f is an MA(2) we examine if the regressions have the zero adjusted r-squares when regressing the filtered series Δs_t^f on the lags greater than 2. Panel B of

[22]We use EViews 6.0 for all econometric tests employed throughout the chapter.

Table 3.3 demonstrates that the conditional heteroscedastic error ζ is able to explain approximately 50% of the variability in $\triangle s_t^f$ on the first two lags for autocorrelations while the adjusted r-squares are zero on those lags greater than 2 (from the 3rd to the 10th). The time series of the filtered spot returns Δs_t^f is indeed an MA(2).

In Table 3.4, we report that the time series of spot returns Δs_t is an ARMA(2,2) in both the baseline and sensitivity analysis. In Panel A of Table 3.4, almost all parameter estimates on the AR(1), AR(2), MA(1), and MA(2) terms of the ARMA(2,2) model for Δs_t, except those in $\gamma = 0.7$, are significant while all constants are not significant. We use the Ljung-Box test for autocorrelations and partial autocorrelations of the equation residuals. In Panel B of Table 3.4, the Ljung-Box Q-statistics and their p-values of the correlogram at each lag are highly significant up to 36 lags of the specified order. The Ljung-Box test statistic rejects the null hypothesis of no autocorrelation at the 1% significance level for almost all lags except the lag 5 in $\gamma = 0.7$. The time series of the spot returns Δs_t is indeed an ARMA(2,2).

3.6.2 Conditional heteroscedasticity

The identification of conditional heteroscedasticity is often determined by testing whether squared or absolute returns are autocorrelated[23]. We use the ARCH-LM test[24] to test if the simulated spot return itself (without an ARMA process) and the simulated spot return of an ARMA(2,2) process from which the innovation $(\widehat{\zeta_t})$ in Equation (3.29) comes are conditionally heteroscedastic. In detail, we regress the squared residuals of the simulated spot return itself and the squared innovations of the ARMA(2,2) spot returns on a constant and lagged squared residuals up to order 9, respectively. We quote the value and significance of the test. Table 3.5 shows that both the F-statistic

[23]See Rodríguez and Ruiz (2003) for the details.

[24]The ARCH test is a Lagrange multiplier (LM) test for autoregressive conditional heteroskedasticity (ARCH) in the residuals (Engle 1982). The ARCH test is frequently applied to raw returns data.

$\Delta s_t = b_0 + b_1 \Delta s_{t-1} + b_2 \Delta s_{t-2} + b_3 \zeta_{t-1} + b_4 \zeta_{t-2}$

		Sensitivity analysis			
	Baseline	$\gamma = 0.7$	$\delta = -0.0025$	$\rho(\pi) = 0$	$\phi = 0.995$
Panel A: parameter estimates					
b_0	1.88E-05	-0.0002	0.0006	-2.84E-05	-5.25E-05
	(0.03)	(-0.22)	(0.72)	(-0.04)	(-0.06)
b_1	-0.407**	0.236	0.310*	-0.526*	-0.856*
	(-2.34)	(0.47)	(42.27)	(-12.49)	(-16.02)
b_2	0.065*	-0.081	-0.977*	0.424*	-0.800*
	(2.62)	(-0.57)	(-135.41)	(10.09)	(-15.46)
b_3	0.426*	-0.241	-0.322*	0.619*	0.854*
	(2.45)	(-0.48)	(-35.999)	(13.88)	(14.80)
b_4	-0.092*	0.098	0.964*	-0.299*	0.760*
	(-3.54)	(0.68)	(108.94)	(-6.74)	(13.51)
Panel B: autocorrelation function - Ljung-Box Q-statistics					
Lag 5	21.92*	1.25	31.58*	39.29*	12.24*
Lag 6	23.10*	17.42*	31.84*	52.32*	27.26*
Lag 7	43.50*	34.13*	33.80*	79.02*	35.45*
Lag 8	94.06*	35.04*	34.34*	86.73*	35.63*
Lag 9	94.09*	35.64*	36.81*	87.00*	42.92*
Lag 10	96.27*	62.53*	38.01*	87.03*	45.35*
Lag 11	125.95*	63.84*	41.10*	108.99*	58.48*
Lag 12	128.56*	65.08*	41.10*	125.11*	60.56*
Lag 13	128.74*	65.16*	52.52*	126.48*	60.57*
Lag 14	131.16*	80.14*	54.17*	152.42*	69.87*
Lag 15	170.39*	84.86*	60.11*	156.55*	98.98*
Lag 16	172.84*	85.65*	60.25*	156.77*	102.52*
Lag 17	175.52*	85.66*	68.44*	168.16*	106.75*
Lag 18	178.91*	89.22*	97.16*	203.27*	106.77*
Lag 19	180.06*	90.18*	100.20*	210.43*	143.29*
Lag 20	203.46*	90.19*	105.38*	229.71*	144.66*
Lag 21	208.70*	108.18*	107.90*	231.47*	146.22*
Lag 22	211.01*	120.63*	129.35*	231.98*	171.55*
Lag 23	224.30*	124.49*	135.93*	272.40*	183.62*
Lag 24	250.72*	124.50*	136.31*	289.97*	183.70*
Lag 25	276.47*	127.92*	136.40*	317.54*	231.46*
Lag 26	282.12*	135.13*	136.53*	330.41*	232.76*
Lag 27	291.42*	154.12*	142.92*	332.55*	237.85*
Lag 28	306.40*	156.72*	163.95*	338.43*	247.68*
Lag 29	307.41*	158.20*	164.29*	339.59*	247.70*
Lag 30	308.27*	159.22*	168.17*	357.51*	261.98*
Lag 31	347.08*	159.97*	168.27*	386.10*	270.27*
Lag 32	379.88*	166.66*	169.24*	419.37*	287.26*
Lag 33	380.04*	170.14*	189.73*	420.40*	287.27*
Lag 34	425.18*	181.09*	190.07*	423.04*	350.75*
Lag 35	431.59*	181.20*	191.49*	453.34*	351.61*
Lag 36	431.62*	181.88*	192.77*	461.11*	355.35*

Notes: t-ratios in parentheses; * and ** denote significance at the 1% and 5% levels, respectively; Q-statistic probabilities are adjusted for 4 ARMA term(s).

Table 3.4: The theoretical spot returns of an ARMA(2,2) process

and χ^2-statistic of the test in Panel A and B are very significant (at 1%) in both the baseline and sensitivity analysis, respectively, suggesting the presence of ARCH in the simulated spot returns. The simulated data for spot returns and for the innovation in the spot returns process are definitely conditionally heteroscedastic.

ARCH-LM heteroscedasticity test					
		Sensitivity analysis			
	Baseline	$\gamma = 0.7$	$\delta = -0.0025$	$\rho(\pi) = 0$	$\phi = 0.995$
Panel A: the simulated spot returns					
F-statistic	375.289*	362.624*	416.187*	513.570*	298.724*
χ^2-statistic	2526.07*	2461.83*	2726.23*	3162.22*	2120.08*
Panel B: the simulated spot returns of an ARMA (2,2)					
F-statistic	372.800*	362.008*	424.955*	417.074*	293.176*
χ^2-statistic	2513.41*	2458.55*	2767.60*	2730.30*	2088.85*

Notes: * denotes significance at the 1% level.

Table 3.5: Testing conditional heteroscedasticity for the theoretical spot returns

We have numerically solved Moore and Roche's model (Equation ([A10],3.19)) and simulated the artificial data, which have the same time series properties as both those found in Moore and Roche (2006) and those implied by the theoretical model. We have also assessed the sensitivity of the results to parameter changes. The theoretical model is able to explain conditional volatility of exchange rates. We find ARCH effects in the simulated spot returns, where the simulated data for spot return and for innovation in the spot return process are definitely conditionally heteroscedastic. We will treat the simulated data and its derivatives as though they were real world data and apply GARCH class models directly to their innovations[25] ($\widehat{\zeta_t}$) that are subject to conditional volatility. In the what follows, we are going to establish the exact dynamic form conditionally heteroscedasticity takes and see whether or not the dynamics match those from the actual monthly data, as in Chapter 2.

[25]Innovations $\widehat{\zeta_t}$ in Equation (3.29) are available after estimating Δs_t^f of an MA(2) in Equation (3.28) using the simulated S in Equation ([A10],3.19).

3.7 The form of conditional heteroscedasticity

We estimate the form of the conditional heteroscedasticity implied in Moore and Roche's model. Generalized autoregressive conditionally heteroscedastic (GARCH) class models are used to capture conditional volatility. Specifically, we employ symmetric (ARCH, GARCH and GARCH-M[26]) and asymmetric (TARCH/GJR, EGARCH, PARCH and CGARCH) conditional volatility models for the best estimates. We present impulse response function (IRF) for the best GARCH models for the ARCH processes we estimate. We then draw the IRFs to establish the exact dynamic form the conditional heteroscedasticity takes. After comparing the IRFs, we show that the two IRFs from the simulated quarterly and from the empirical monthly data "look the same", with approximately monotonic decreasing profiles. We conclude that the Lucas two-country monetary model with habit is capable of producing the same kind of ARCH features as we see in the real data. For simplicity, we call the Lucas two-country monetary model with habit in Moore and Roche (2006) "the theoretical model", the GARCH class conditional volatility models "the empirical model", the data simulated by using the theoretical model "the theoretical data" and the spot USD/GBP exchange rates collected from Thomson Datastream "the empirical data". To be more specific, the empirical spot returns consist of the empirical monthly spot return[27], which is obtained from monthly averages of daily spot rates in that month, and the empirical quarterly

[26]With the same conditional variance equation as in the GARCH model, the GARCH-in-mean (GARCH-M for short) model has a different conditional mean equation where the conditional variance of asset returns enters into the conditional mean equation, for example, $y_t = c + bx_t + dh_t^2 + \tau_t$ where $\tau_t \sim N(0, h_t^2)$, which says that the return is partly determined by its risk. The GARCH-M model is often used in financial applications where the expected return on an asset is related to the expected asset risk. The estimated coefficient on the expected risk is a measure of the risk-return tradeoff. In empirical applications, the conditional variance term, h_t^2, appears directly in the conditional mean equation, rather than in square root form h_t (p480, Brooks (2002)). See Section 2.4 in Chapter 2 for the details of the other GARCH class models.

[27]We use the same empirical monthly data set as those employed in Chapter 2. The time series of the empirical monthly spot return is constructed by using the daily spot USD/GBP exchange rate, where the average of daily prices in that month (quarter) as the proxy of monthly (quarterly) price. The empirical monthly data spans the period from 1973 to 2005, which is the same as in Moore and Roche (2006) in terms of the spot GBP/USD exchange rate. See Section 2.3 in Chapter 2 for the details.

spot return which obtained from quarterly point spot rates[28] in time.

3.7.1 Estimation of GARCH models

We estimate the best fit GARCH model(s) for the time series of the theoretical (simulated) spot return Δs_t and for the time series of the empirical (real) monthly and quarterly USD/GBP spot returns, respectively, in terms of significance of coefficients, asymmetric effects and persistent shocks as well as the relationship of return with risk. We then do the same for the innovation $\widehat{\zeta}_t$ in the time series representation for Δs_t that is an ARMA(2,2) and the residuals in the MA(1) process for the empirical monthly spot return and in the ARMA(2,3) process for the empirical quarterly spot return, respectively. The first exercise is a misspecification if Moore and Roche's model is "true (i.e. under the null of Moore and Roche (2006)). But the reason for considering the FOREX spot return itself in the first exercise is simply because this is exactly what empirical researchers tend to do. Also we could argue that the persistence properties in the FOREX spot return at the quarterly frequency are rather weak, so that modelling ARCH effects in Δs itself rather than the innovation in the time series model for Δs is a minor misspecification. This exercise also ties in closely with the results in Chapter 2 in which we find the best fit forecasting model for the spot return rather than its innovation.

3.7.1.1 Theoretical estimation

We found the presence of ARCH effects in the theoretical spot return Δs_t using the Engle (1982) test, as we did in Section 3.6. It suggests that the GARCH class models are appropriate for the theoretical data. In Table 3.6 and Table 3.7, we report the

[28] The empirical quarterly point spot rates of USD/GBP in time cover the sample period from 1973:1 to 2005:4, which is the same as in Moore and Roche (2006) in terms of the spot GBP/USD exchange rate.

value and significance of the estimates of the GARCH(1,1)[29] class conditional volatility models for the theoretical data in baseline.

We report the results of estimating GARCH class models for the simulated Δs_t series itself in baseline in Table 3.6. For all specifications of the GARCH class models we employ, the coefficients on the lagged squared error (ARCH) term in the conditional variance equation are statistically significant, except for the non-significant ARCH term in the CGARCH(1,1) model. Meanwhile, the coefficients on the lagged conditional variance (GARCH) term are statistically significant, except for the one in the EGARCH(1,1) model. The asymmetry terms (η, μ, v and γ) in the EGARCH, TARCH, PARCH and CGARCH models respectively are not significant. The sum of the ARCH and GARCH coefficients for the GARCH, GARCH-M, EGARCH and PARCH models respectively is (approximately) close to unity, which implies that shocks to conditional variance will be (highly) persistent. This can be found by using the models to forecast future values of the conditional variance for the real USD/GBP spot returns, as in Chapter 2. A large sum of these coefficients (e.g. the TARCH model) will imply that a large positive or a large negative return will lead future forecasts of the variance to be high for a protracted period. The variance intercept terms (ω) are significant except the non-significant one in the CGARCH model, where the variance intercept terms (ω) in the ARCH, GARCH, GARCH-M and TARCH models are very small, while the coefficients on the significant GARCH terms are larger ($\gtrsim 0.4$). The conditional standard deviation term that is introduced into the mean equation of the GARCH-M model is not significant, which suggests that the property that higher market-wide risk would lead to higher returns is not available. We report the results of additional ARCH effects

[29]We found that the GARCH(1,1) class models have better estimating performances than the GARCH(p,q) models with higher orders ($1 < p \leqq 9$, $1 < q \leqq 9$) when we estimated the GARCH(9,9) models then removed insignificant lags one at a time (re-estimating each time).

Mean Equation

$$\Delta s_t = c + \tau_t \quad \tau_t \backsim N(0, h_t^2)$$

(where $\Delta s_t = c + dh_t + \tau_t \quad \tau_t \sim N(0, h_t^2)$ to GARCH(1,1)-M)

Variance Equation

$h_t^2 = \omega + \alpha\tau_{t-1}^2$	ARCH(1)
$h_t^2 = \omega + \alpha\tau_{t-1}^2 + \beta h_{t-1}^2$	GARCH(1,1)
$h_t^2 = \omega + \alpha\tau_{t-1}^2 + \beta h_{t-1}^2$	GARCH(1,1)-M
$\ln(h_t^2) = \omega + \alpha\left[\left\vert\frac{\tau_{t-1}}{h_{t-1}}\right\vert - \sqrt{\frac{2}{\pi}}\right] + \beta\ln(h_{t-1}^2) + \eta\frac{\tau_{t-1}}{\vert h_{t-1}\vert}$	EGARCH(1,1)
$h_t^2 = \omega + \alpha\tau_{t-1}^2 + \beta h_{t-1}^2 + \mu\tau_{t-1}^2 I_{t-1}$ where $I_{t-1} = 1$ if $\tau_{t-1} < 0$ and 0 otherwise	TARCH(1,1)
$h_t^\vartheta = \omega + \alpha\left(\vert\tau_{t-1}\vert - \upsilon\tau_{t-1}\right)^\vartheta + \beta h_{t-1}^\vartheta$	PARCH(1,1)
$h_t^2 = \omega + \rho\left(q_{t-1} - \omega\right) + \varphi\left(\tau_{t-1}^2 - h_{t-1}^2\right) + \alpha\left(\tau_{t-1}^2 - q_{t-1}\right) + \gamma\left(\tau_{t-1}^2 - q_{t-1}\right)D_{t-1} + \beta\left(h_{t-1}^2 - q_{t-1}\right)$	CGARCH(1,1)

Estimating model	c	d	ω	α	β	η	μ	υ	ϑ	ρ	φ	γ
ARCH(1)	0.00049	-	0.0000912*	15.077*	-	-	-	-	-	-	-	-
	(0.74)	-	(4.29)	(7.15)	-	-	-	-	-	-	-	-
GARCH(1,1)	0.00016*	-	0.0000027*	0.578*	0.420*	-	-	-	-	-	-	-
	(3.46)	-	(7.99)	(15.80)	(11.48)	-	-	-	-	-	-	-
GARCH(1,1)-M	0.00048*	-0.045	0.0000023*	0.447*	0.551*	-	-	-	-	-	-	-
	(5.23)	(-1.37)	(10.09)	(23.59)	(29.05)	-	-	-	-	-	-	-
EGARCH(1,1)	0.000006	-	-6.054*	1.017***	0.164	-0.016	-	-	-	-	-	-
	(0.01)	-	(-3.31)	(1.95)	(0.64)	(-0.17)	-	-	-	-	-	-
TARCH(1,1)	0.00025*	-	0.0000029*	1.213*	0.398*	-	0.193	-	-	-	-	-
	(4.80)	-	(675.74)	(11.76)	(21.99)	-	(1.18)	-	-	-	-	-
PARCH(1,1)	0.00029*	-	0.005**	0.537*	0.605*	-	-	-0.019	0.644*	-	-	-
	(6.08)	-	(2.39)	(12.27)	(27.96)	-	-	(-0.38)	(7.47)	-	-	-
CGARCH(1,1)	0.00026*	-	0.002	-0.038	0.514**	-	-	-	-	0.999*	0.445*	0.030
	(5.67)	-	(0.04)	(-1.03)	(2.28)	-	-	-	-	(32.32)	(11.82)	(0.76)

Notes: z-statistics in parentheses; *, ** and *** denote significance at the 1%, 5% and 10% levels, respectively.

Table 3.6: Estimates of GARCH class conditional volatility models for the theoretical quarterly spot return itself in baseline

up to the order 9 in the residuals after the model estimates in Appendix B.2. The presence of additional ARCH is in the residuals of the estimated models, except with the PARCH and CGARCH models[30].

Next, we report the estimating results of the GARCH class models for the innovation for Δs_t that is an ARMA(2,2) in the baseline in Table 3.7. For all cases, the coefficient estimates on the variance intercept, ARCH and GARCH terms in the conditional variance equation are highly statistically significant (at 1%). Also, the coefficient estimates on the asymmetry terms (η and γ) in the EGARCH and CGARCH models are highly statistically significant (at 1%), which suggests, as expected, that negative shocks imply a higher next period conditional variance than positive shocks of the same magnitude. The persistence of volatility shocks is found with the GARCH, GARCH-M, and PARCH models due to the sum of the ARCH and GARCH coefficients close to unity; compared with a large sum ($\gtrsim 1.6$) of these coefficients for the EGARCH and TARCH models and a small sum (≈ 0.7) for the CGARCH model. We find the presence of the additional ARCH effects in the residuals for the EGARCH and CGARCH models[31]. Again, the conditional standard deviation term in the mean equation of the GARCH-M model is not significant.

It is noted that EGARCH is the best fit model to additional ARCH effects due to the presence of ARCH effects in the residuals for both the simulated Δs_t and its innovation. EGARCH and CGARCH are the best fit models to asymmetry effects because of the significant asymmetric terms in the conditional variance equation as shown in Table 3.7. PARCH is the best fit asymmetric model to persistent shocks in terms of the sum of the ARCH and GARCH coefficients close to unity as found in both Table 3.6 and Table 3.7. Taking into account the properties of conditional volatility, which is not only

[30] See Appendix B.2.1 for the details.
[31] See Appendix B.2.2 for the details.

Mean Equation

$$\Delta s_t = c + c_1 \Delta s_{t-1} + c_2 \Delta s_{t-2} + c_3 \tau_{t-1} + c_4 \tau_{t-2} + \tau_t \quad \tau_t \backsim N(0, h_t^2)$$

(where $\Delta s_t = c + c_1 \Delta s_{t-1} + c_2 \Delta s_{t-2} + c_3 \tau_{t-1} + c_4 \tau_{t-2} + dh_t + \tau_t \quad \tau_t \sim N(0, h_t^2)$ to GARCH(1,1)-M)

Variance Equation

$$h_t^2 = \omega + \alpha \tau_{t-1}^2$$ ARCH(1)

$$h_t^2 = \omega + \alpha \tau_{t-1}^2 + \beta h_{t-1}^2$$ GARCH(1,1)

$$h_t^2 = \omega + \alpha \tau_{t-1}^2 + \beta h_{t-1}^2$$ GARCH(1,1)-M

$$\ln(h_t^2) = \omega + \alpha \left[\left| \frac{\tau_{t-1}}{h_{t-1}} \right| - \sqrt{\frac{2}{\pi}} \right] + \beta \ln(h_{t-1}^2) + \eta \frac{\tau_{t-1}}{|h_{t-1}|}$$ EGARCH(1,1)

$$h_t^2 = \omega + \alpha \tau_{t-1}^2 + \beta h_{t-1}^2 + \mu \tau_{t-1}^2 I_{t-1} \text{ where } I_{t-1} = 1 \text{ if } \tau_{t-1} < 0 \text{ and 0 otherwise}$$ TARCH(1,1)

$$h_t^\vartheta = \omega + \alpha \left(|\tau_{t-1}| - \upsilon \tau_{t-1} \right)^\vartheta + \beta h_{t-1}^\vartheta$$ PARCH(1,1)

$$h_t^2 = \omega + \rho (q_{t-1} - \omega) + \varphi \left(\tau_{t-1}^2 - h_{t-1}^2 \right) + \alpha \left(\tau_{t-1}^2 - q_{t-1} \right) + \gamma \left(\tau_{t-1}^2 - q_{t-1} \right) D_{t-1} + \beta \left(h_{t-1}^2 - q_{t-1} \right)$$ CGARCH(1,1)

Estimating model	c	c_1	c_2	c_3	c_4	d	ω	α	β	η	μ	υ	ϑ	ρ	φ	γ
ARMA(2,2)-ARCH(1)	-0.00026	-0.825*	0.015	0.903*	0.032	-	0.0000117*	38.372*	-	-	-	-	-	-	-	-
	(-1.07)	(-2.81)	(0.16)	(2.93)	(0.21)	-	(244.33)	(3.43)	-	-	-	-	-	-	-	-
ARMA(2,2)-GARCH(1,1)	0.00017*	1.203*	-0.218*	-1.401*	0.412*	-	0.0000025*	0.463*	0.535*	-	-	-	-	-	-	-
	(5.30)	(21.25)	(-4.19)	(-24.66)	(7.77)	-	(10.08)	(22.49)	(25.92)	-	-	-	-	-	-	-
ARMA(2,2)-GARCH(1,1)-M	-0.00039	1.382*	-0.535*	-1.319*	0.531*	-0.049	0.0000037*	0.711*	0.287*	-	-	-	-	-	-	-
	(-0.29)	(12.66)	(-6.30)	(-13.23)	(7.21)	(-0.63)	(5.89)	(7.99)	(3.22)	-	-	-	-	-	-	-
ARMA(2,2)-EGARCH(1,1)	0.00073*	0.998*	-0.103***	-1.121*	0.218*	-	-1.099*	0.948*	0.943*	-0.049*	-	-	-	-	-	-
	(11.57)	(11.13)	(-1.66)	(-12.84)	(3.53)	-	(-20.56)	(38.36)	(182.28)	(-3.43)	-	-	-	-	-	-
ARMA(2,2)-TARCH(1,1)	0.00019*	-0.303	0.068**	0.173	-0.171**	-	0.0000028*	1.171*	0.401*	-	0.186	-	-	-	-	-
	(6.45)	(-0.80)	(2.31)	(0.45)	(-2.12)	-	(9.93)	(12.96)	(22.13)	-	(1.31)	-	-	-	-	-
ARMA(2,2)-PARCH(1,1)	-0.0000048	-0.735*	0.112*	0.730*	-0.116*	-	0.004*	0.526*	0.617*	-	-	-0.026	0.744*	-	-	-
	(-0.13)	(-8.29)	(3.14)	(8.85)	(-2.72)	-	(2.79)	(13.41)	(29.04)	-	-	(-0.55)	(10.70)	-	-	-
ARMA(2,2)-CGARCH(1,1)	-0.004*	0.889*	0.115	-0.869*	-0.093	-	0.000068*	0.097*	0.623*	-	-	-	-	0.957*	0.152*	0.234*
	(-6.22)	(12.13)	(1.57)	(-12.18)	(-1.34)	-	(49467.35)	(5.40)	(32.80)	-	-	-	-	(256.20)	(18.48)	(9.30)

Notes: z-statistics in parentheses; *, ** and *** denote significance at the 1%, 5% and 10% levels, respectively.

Table 3.7: Estimates of GARCH class conditional volatility models for the theoretical quarterly spot return of an ARMA(2,2) in baseline

conditional but also asymmetric, the asymmetric CGARCH, EGARCH and PARCH conditional volatility models are the best fit estimating models for the simulated data. Further details of the selection process are provided in Appendix B.2.

We also look at the cases apart from the baseline and assess sensitivity to parameter changes. We report the results of the significant coefficients, asymmetry effects, persistent shocks, additional ARCH, and non-significant conditional variance terms in the mean equation of the GARCH-M models in the sensitivity analysis in Appendix B.2, which is consistent with what we find in the baseline.

The results of estimates and additional ARCH for the other models[32] and for steps in model selection process are given in Appendix B.2. For other cases, we find not only consistent results but also more information compared to that disclosed in Table 3.6 and Table 3.7. For example, the conditional variance term introduced into the mean equation for the GARCH-M model has a positive sign and is highly significant (at 1%). This suggests that higher market-wide risk, proxied by the conditional variance, will lead to higher returns. Thus the parameter (d) of the conditional variance in the mean equation of the GARCH-M model can be interpreted as a risk premium. The theoretical model is able to capture the relationship between return and risk where return is partly determined by risk.

We summarize the findings in both the baseline and sensitivity analysis, as well as the results for the other relevant models in detail in Appendix B.2. Based on the analysis and summary, the main conclusions are as follows:

- GARCH class conditional volatility models are appropriate for the theoretical quarterly FOREX data in which the predictable properties of conditional volatil-

[32]We report the best fit GARCH estimates for the simulated filtered series Δs_t^f shown in Equation (3.28) and the simulated genuine innovation $\widehat{\zeta}_t$ shown in Equation (3.29) in Appendix B.2.3 and B.2.4. Our aim is to maximally capture ARCH basing on the model properties even its implied properties.

ity are found.

- The presence of additional ARCH effects is in the residuals of the estimated standard GARCH class models.

- In symmetric conditional volatility models, GARCH(1,1) is the best fit estimating model to conditional volatility for the theoretical spot return and its innovation.

- In asymmetric conditional volatility models, CGARCH, EGARCH and PARCH are the best fit estimating models to conditional volatility for the theoretical spot return and its innovation in terms of the properties of asymmetry, additional ARCH, and persistent volatility shock, respectively.

- In the sensitivity analysis, more significant results on ARCH effects, asymmetric effects and persistent volatility shocks to conditional volatility are found than those disclosed in the baseline. $\gamma = 0.7$ is the best performer in the sensitivity analysis.

- The theoretical model can generate realistic conditional volatility even asymmetry conditional volatility.

As stated above, the consistent estimating results suggest that the asymmetric CGARCH, EGARCH and PARCH conditional volatility models are the best fit GARCH models to conditional volatility for the theoretical data.

3.7.1.2 Empirical estimation

We turn to focus on the actual data. As we know, the GARCH class conditional volatility models are appropriate for the empirical monthly spot return[33] of the USD/GBP exchange rate. We report the model estimates for the monthly USD/GBP spot return itself in Table 3.8. For all specifications, the coefficients on the lagged squared error (ARCH) term in the conditional variance equation are highly statistically significant (at 1%) except the non-significant ARCH term in the CGARCH(1,1) model. All coefficients on the lagged conditional variance (GARCH) term are highly statistically significant (at 1%). The asymmetry term (γ) in the conditional variance equation of the CGARCH model is very significant with a positive sign, suggesting that negative shocks imply a higher next period conditional variance than positive shocks of the same magnitude. The sum of the coefficients on the ARCH and GARCH terms of the PARCH model compared with that of these two coefficients of the other models is approximately close to unity (≈ 0.92), which implies that shocks to conditional variance will be persistent. The variance intercept terms (ω) are significant except the non-significant one in the PARCH model, where the variance intercept terms (ω) in ARCH, GARCH, GARCH-M and TARCH are very small. It is found that the conditional variance term in the mean equation of the GARCH-M model is not significant. The presence of additional ARCH is in the residuals of the estimated ARCH model, while there is no additional ARCH in the residuals for other estimated GARCH class models.

Moreover, we report the estimating results of the GARCH class models for the

[33] In Chapter 2, we find that the time series of the monthly USD/GBP spot return is stationary using a unit root test (ADF test). We compute the Engle (1982) test for ARCH effects to make sure that the GARCH-type models are appropriate for the data. We find the highly significant (at 1%) F-statistic and LM-statistic of the test by regressing the squared residuals on a constant and 9 lags. The presence of ARCH is in the residuals for the monthly USD/GBP return.

Mean Equation

$$\Delta s_t = c + \tau_t \quad \tau_t \backsim N(0, h_t^2)$$

(where $\Delta s_t = c + dh_t^2 + \tau_t \quad \tau_t \sim N(0, h_t^2)$ to GARCH(1,1)-M)

Variance Equation

$$h_t^2 = \omega + \alpha\tau_{t-1}^2 \qquad \text{ARCH(1)}$$

$$h_t^2 = \omega + \alpha\tau_{t-1}^2 + \beta h_{t-1}^2 \qquad \text{GARCH(1,1)}$$

$$h_t^2 = \omega + \alpha\tau_{t-1}^2 + \beta h_{t-1}^2 \qquad \text{GARCH(1,1)-M}$$

$$\ln(h_t^2) = \omega + \alpha\left[\left|\frac{\tau_{t-1}}{h_{t-1}}\right| - \sqrt{\frac{2}{\pi}}\right] + \beta\ln(h_{t-1}^2) + \eta\frac{\tau_{t-1}}{|h_{t-1}|} \qquad \text{EGARCH(1,1)}$$

$$h_t^2 = \omega + \alpha\tau_{t-1}^2 + \beta h_{t-1}^2 + \mu\tau_{t-1}^2 I_{t-1} \text{ where } I_{t-1} = 1 \text{ if } \tau_{t-1} < 0 \text{ and 0 otherwise} \qquad \text{TARCH(1,1)}$$

$$h_t^\vartheta = \omega + \alpha\left(|\tau_{t-1}| - \upsilon\tau_{t-1}\right)^\vartheta + \beta h_{t-1}^\vartheta \qquad \text{PARCH(1,1)}$$

$$h_t^2 = \omega + \rho\left(q_{t-1} - \omega\right) + \varphi\left(\tau_{t-1}^2 - h_{t-1}^2\right) + \alpha\left(\tau_{t-1}^2 - q_{t-1}\right) + \gamma\left(\tau_{t-1}^2 - q_{t-1}\right)D_{t-1} + \beta\left(h_{t-1}^2 - q_{t-1}\right) \qquad \text{CGARCH(1,1)}$$

Estimating model	c	d	ω	α	β	η	μ	υ	ϑ	ρ	φ	γ
ARCH(1)	-0.001	-	0.00042*	0.322*	-	-	-	-	-	-	-	-
	(-1.19)	-	(9.46)	(3.03)	-	-	-	-	-	-	-	-
GARCH(1,1)	-0.00057	-	0.00009**	0.259*	0.609*	-	-	-	-	-	-	-
	(-0.62)	-	(2.48)	(3.26)	(5.74)	-	-	-	-	-	-	-
GARCH(1,1)-M	0.00089	-2.609	0.00009**	0.257*	0.607*	-	-	-	-	-	-	-
	(0.48)	(-0.79)	(2.46)	(3.19)	(5.58)	-	-	-	-	-	-	-
EGARCH(1,1)	-0.00040	-	-1.384**	0.392*	0.857*	-0.015	-	-	-	-	-	-
	(-0.38)	-	(-2.49)	(4.31)	(12.44)	(-0.31)	-	-	-	-	-	-
TARCH(1,1)	-0.00063	-	0.00009**	0.215*	0.611*	-	0.065	-	-	-	-	-
	(-0.60)	-	(2.48)	(2.72)	(5.66)	-	(0.64)	-	-	-	-	-
PARCH(1,1)	-0.00023	-	0.007	0.190*	0.732*	-	-	0.005	0.773***	-	-	-
	(-0.22)	-	(0.70)	(3.36)	(8.55)	-	-	(0.03)	(1.94)	-	-	-
CGARCH(1,1)	0.00081	-	0.00060*	-0.072	0.701*	-	-	-	-	0.938*	0.089	0.293*
	(0.92)	-	(3.92)	(-0.75)	(6.44)	-	-	-	-	(19.18)	(1.16)	(2.81)

Notes: z-statistics in parentheses; *, ** and *** denote significance at the 1%, 5% and 10% levels, respectively.

The time series of the empirical monthly spot return itself is obtained from monthly averages of daily spot rates in that month.

Table 3.8: Estimates of GARCH class conditional volatility models for the empirical monthly spot return of USD/GBP itself

Mean Equation

$$\Delta s_t = c + c_1\tau_{t-1} + \tau_t \quad \tau_t \backsim N(0, h_t^2)$$

(where $\Delta s_t = c + c_1\tau_{t-1} + dh_t^2 + \tau_t \quad \tau_t \sim N(0, h_t^2)$ to GARCH(1,1)-M)

Variance Equation

$h_t^2 = \omega + \alpha\tau_{t-1}^2$	ARCH(1)
$h_t^2 = \omega + \alpha\tau_{t-1}^2 + \beta h_{t-1}^2$	GARCH(1,1)
$h_t^2 = \omega + \alpha\tau_{t-1}^2 + \beta h_{t-1}^2$	GARCH(1,1)-M
$\ln(h_t^2) = \omega + \alpha\left[\left\|\frac{\tau_{t-1}}{h_{t-1}}\right\| - \sqrt{\frac{2}{\pi}}\right] + \beta\ln(h_{t-1}^2) + \eta\frac{\tau_{t-1}}{\|h_{t-1}\|}$	EGARCH(1,1)
$h_t^2 = \omega + \alpha\tau_{t-1}^2 + \beta h_{t-1}^2 + \mu\tau_{t-1}^2 I_{t-1}$ where $I_{t-1} = 1$ if $\tau_{t-1} < 0$ and 0 otherwise	TARCH(1,1)
$h_t^{\vartheta} = \omega + \alpha\left(\|\tau_{t-1}\| - \upsilon\tau_{t-1}\right)^{\vartheta} + \beta h_{t-1}^{\vartheta}$	PARCH(1,1)
$h_t^2 = \omega + \rho(q_{t-1} - \omega) + \varphi\left(\tau_{t-1}^2 - h_{t-1}^2\right) + \alpha\left(\tau_{t-1}^2 - q_{t-1}\right) + \gamma\left(\tau_{t-1}^2 - q_{t-1}\right)D_{t-1} + \beta\left(h_{t-1}^2 - q_{t-1}\right)$	CGARCH(1,1)

Estimating model	c	c_1	d	ω	α	β	η	μ	υ	ϑ	ρ	φ	γ
MA(1)-ARCH(1)	-0.00082	0.410*	-	0.00042*	0.187**	-	-	-	-	-	-	-	-
	(-0.53)	(7.76)	-	(9.07)	(2.04)	-	-	-	-	-	-	-	-
MA(1)-GARCH(1,1)	-0.00078	0.396*	-	0.00006***	0.139*	0.754*	-	-	-	-	-	-	-
	(-0.55)	(8.43)	-	(1.90)	(2.73)	(8.31)	-	-	-	-	-	-	-
MA(1)-GARCH(1,1)-M	0.00033	0.396*	-2.674	0.00006***	0.140*	0.752*	-	-	-	-	-	-	-
	(0.12)	(8.41)	(-0.43)	(1.88)	(2.70)	(8.13)	-	-	-	-	-	-	-
MA(1)-EGARCH(1,1)	-0.00048	0.405*	-	-12.511*	0.226***	-0.622*	-0.002	-	-	-	-	-	-
	(-0.30)	(8.48)	-	(-7.98)	(1.75)	(-3.01)	(-0.03)	-	-	-	-	-	-
MA(1)-TARCH(1,1)	-0.00084	0.396*	-	0.00006***	0.130**	0.754*	-	0.014	-	-	-	-	-
	(-0.57)	(8.38)	-	(1.93)	(2.45)	(8.53)	-	(0.19)	-	-	-	-	-
MA(1)-PARCH(1,1)	-0.00090	0.390*	-	0.00069	0.142*	0.772*	-	-	-0.003	1.344	-	-	-
	(-0.61)	(8.39)	-	(0.32)	(3.16)	(8.77)	-	-	(-0.02)	(1.58)	-	-	-
MA(1)-CGARCH(1,1)	-0.00067	0.385*	-	0.00053*	-0.030	-0.345	-	-	-	-	0.912*	0.107*	0.139
	(-0.48)	(8.11)	-	(4.91)	(-0.40)	(-0.57)	-	-	-	-	(17.02)	(2.80)	(1.08)

Notes: z-statistics in parentheses; *, ** and *** denote significance at the 1%, 5% and 10% levels, respectively.

The time series of the empirical monthly MA(1) spot return is obtained from monthly averages of daily spot rates in that month.

Table 3.9: Estimates of GARCH class conditional volatility models for the empirical monthly MA(1) spot return of USD/GBP

innovation of the monthly USD/GBP spot return that is of an MA(1) process[34] in Table 3.9.

For all specifications in Table 3.9, the coefficients on the ARCH, and GARCH terms in the conditional variance equation are statistically significant, except in the CGARCH(1,1) model, where neither are. Particularly, all GARCH coefficient estimates except the one in the CGARCH model are highly statistically significant (at 1%). None of the asymmetry terms in the asymmetric models is significant. Also in most cases, the variance intercept terms (ω) are significant except the non-significant intercept term in the PARCH model. The variance intercept terms in ARCH, GARCH, GARCH-M and TARCH are very small. Again, the PARCH model is able to capture persistent volatility shocks due to the sum of its ARCH and GARCH terms approximately close to unity (≈ 0.91) and the conditional variance term in the mean equation of the GARCH-M model is not significant. We find the presence of ARCH effects up to the order 9 in the residuals of the estimated ARCH and EGARCH models using the Engle (1982) test.

So far, we have estimated the best fit GARCH models for the spot returns of the theoretical and empirical data and then their innovations that are in the time series representations for the corresponding spot returns of an ARMA process. We compare the results in Tables 3.6-3.9. On the one hand, the main results in Table 3.6 and Table 3.8 for the spot return itself (without the ARMA process) are that: 1) in both Table 3.6 and Table 3.8 only the CGARCH model has the non-significant ARCH coefficient estimates; 2) only the EGARCH model in Table 3.6 has the non-significant GARCH coefficient estimate; 3) only the CGARCH model in Table 3.8 has the significant asym-

[34] In Chapter 2, we find that the time series of the monthly USD/GBP spot return is an MA(1) using Schwarz's (1978) Bayesian information criterion (SBIC) that is recommended by Diebold (2001). We also report the estimates of the ARMA model where the MA(1) term is highly significant (at 1%) and the results of the Ljung-Box test. See Chapter 2 for the further details.

metric coefficient estimate in the variance equation. The CGARCH and EGARCH models in Table 3.8 have better estimating performances than they do in Table 3.6, in terms of the significance of the coefficients on the GARCH and asymmetry. The estimating results in Table 3.8 for the empirical spot return itself, where the coefficients on the important terms in the variance equation are significant, are superior to those in Table 3.6 for the theoretical Δs_t itself. On the other hand, the main differences between Table 3.7 and Table 3.9 for the innovations of the ARMA spot returns are that, in Table 3.9, the CGARCH model has the nonsignificant coefficient estimates on the ARCH, GARCH and asymmetric terms and the EGARCH model has the insignificant asymmetric term, while the coefficients on these terms are highly statistically significant in Table 3.7. The CGARCH and EGARCH models in Table 3.7 have better estimating performances than in Table 3.9 due to the significant ARCH, GARCH and asymmetric terms in the variance equation. The estimating results in Table 3.7 for the innovation in the theoretical Δs_t of an ARMA(2,2) process are superior to those in Table 3.9 for the residual in the MA(1) process for the empirical monthly spot return. It is found in common in Tables 3.6-3.9 that the PARCH model is the only asymmetric conditional volatility model that captures persistent shock to conditional variance, and the conditional variance term in the mean equation of the GARCH-M model is globally insignificant, and the variance intercept terms in ARCH, GARCH, GARCH-M and TARCH compared to those in CGARCH, EGARCH and PARCH are very small.

We also note that the results in Table 3.7 are highly similar to those in Table 3.8 while the results in Table 3.6 are highly similar to those in Table 3.9. Specifically, for both Table 3.7 and Table 3.8, all coefficients on the GARCH term are highly statistically significant and asymmetric effects are present, where the asymmetric coefficients in the CGARCH model in both tables and in the EGARCH model in Table 3.7 are highly statistically significant, while none of asymmetric terms is significant in both Table 3.6 and Table 3.9. The GARCH coefficients in the EGARCH model in Table 3.6 and in the CGARCH model in Table 3.9 are insignificant. This suggests that the estimates of conditional volatility for the innovation in the theoretical quarterly spot return of an ARMA(2,2) process perform as consistently as those for the empirical monthly spot return itself without an ARMA process do. At this moment, both cases make the maximal capture of estimating information of conditional volatility in either theoretical or empirical frames. This, we think, could be one reason why empirical researchers tend to consider the FOREX spot return itself and not its innovation, as mentioned at the section's beginning. It also explains why we chose to estimate and forecast volatility for the FOREX spot return rather than its innovation in Chapter 2.

We estimate but do not report the best fit models for the empirical daily and quarterly spot returns of the USD/GBP exchange rate, where the empirical quarterly spot return is obtained from quarterly averages of daily spot rates in that quarter, as the proxy of quarterly prices. We find that the empirical averaging data at a higher (e.g. daily/monthly) frequency is able to provide highly similar properties of conditional volatility to those implied in the theoretical time point data at a low (e.g. quarterly) frequency. In the theoretical framework, modelling ARCH effects in the innovation in the time series presentation for the theoretical spot return is able to explain realistic conditional volatility, even if the theoretical data is at a low frequency. In the empirical framework, realistic conditional volatility could be well captured by using high(er) frequency empirical data for spot return due to "noise" or "imperfection" in the real world. "Noise" mentioned here could be crises, noisy traders, momentum, psychology, central bank intervention and macroeconomic variables etc. in daily life. "Noise" makes empirical data at the same frequency as theoretical data imprecise, so that information is lost or changed. Hence, information obtained by using the high frequency empirical data is able to match that implied by the low frequency theoretical data. This is why we find the consistent results as in Table 3.7 and Table 3.8. It also seems a reasonable solution to use empirical mixed data with a diverse (e.g. from low to high) frequency to capture conditional volatility found in the theoretical data.

After comparing the empirical monthly averaging data to the theoretical quarterly time point data, we also look at the estimating results for the empirical quarterly spot return, which is constructed by using quarterly USD/GBP point spot rates in time. Using the unit root test and Engle (1982) test, we find that the GARCH class models are appropriate for the time series of the quarterly spot return that is stationary. The estimating results for the empirical quarterly point data are (more or less) similar to the results as found consistent for the theoretical quarterly point data and the empirical monthly averaging data.

We report the model estimates for the quarterly USD/GBP spot return itself in Table 3.10. The coefficients on the lagged squared error (ARCH) term in the conditional variance equation of the GARCH-M, EGARCH and CGARCH models are statistically significant while other ARCH coefficient estimates are not significant. For all specifications, the coefficients on the lagged conditional variance (GARCH) term are highly statistically significant (at 1%) except the non-significant GARCH term in the CGARCH(1,1) model. The conditional variance term that appears in the mean equation of the GARCH-M model has a negative sign and is significant at the 10% level, which suggests that higher market-wide risk, proxied by the conditional variance, will lead to lower returns. All estimating results mentioned above are distinguished from those in Table 3.6 and Table 3.8. The asymmetry term in the conditional variance equation of the CGARCH model is very significant with a positive sign, which is the

Mean Equation

$$\Delta s_t = c + \tau_t \quad \tau_t \backsim N(0, h_t^2)$$

(where $\Delta s_t = c + dh_t^2 + \tau_t \quad \tau_t \sim N(0, h_t^2)$ to GARCH(1,1)-M)

Variance Equation

$$h_t^2 = \omega + \alpha\tau_{t-1}^2 \qquad \text{ARCH(1)}$$

$$h_t^2 = \omega + \alpha\tau_{t-1}^2 + \beta h_{t-1}^2 \qquad \text{GARCH(1,1)}$$

$$h_t^2 = \omega + \alpha\tau_{t-1}^2 + \beta h_{t-1}^2 \qquad \text{GARCH(1,1)-M}$$

$$\ln(h_t^2) = \omega + \alpha\left[\left|\frac{\tau_{t-1}}{h_{t-1}}\right| - \sqrt{\frac{2}{\pi}}\right] + \beta\ln(h_{t-1}^2) + \eta\frac{\tau_{t-1}}{|h_{t-1}|} \qquad \text{EGARCH(1,1)}$$

$$h_t^2 = \omega + \alpha\tau_{t-1}^2 + \beta h_{t-1}^2 + \mu\tau_{t-1}^2 I_{t-1} \text{ where } I_{t-1} = 1 \text{ if } \tau_{t-1} < 0 \text{ and 0 otherwise} \qquad \text{TARCH(1,1)}$$

$$h_t^\vartheta = \omega + \alpha\left(|\tau_{t-1}| - \upsilon\tau_{t-1}\right)^\vartheta + \beta h_{t-1}^\vartheta \qquad \text{PARCH(1,1)}$$

$$h_t^2 = \omega + \rho\left(q_{t-1} - \omega\right) + \varphi\left(\tau_{t-1}^2 - h_{t-1}^2\right) + \alpha\left(\tau_{t-1}^2 - q_{t-1}\right) + \gamma\left(\tau_{t-1}^2 - q_{t-1}\right)D_{t-1} + \beta\left(h_{t-1}^2 - q_{t-1}\right) \qquad \text{CGARCH(1,1)}$$

Estimating model	c	d	ω	α	β	η	μ	υ	ϑ	ρ	φ	γ
ARCH(1)	-0.00009	-	0.002*	0.153	-	-	-	-	-	-	-	-
	(-0.02)	-	(5.56)	(1.23)	-	-	-	-	-	-	-	-
GARCH(1,1)	-0.002	-	0.00016	0.073	0.867*	-	-	-	-	-	-	-
	(-0.36)	-	(0.86)	(1.39)	(7.88)	-	-	-	-	-	-	-
GARCH(1,1)-M	0.039***	-15.74***	0.00067***	0.153***	0.591*	-	-	-	-	-	-	-
	(1.66)	(-1.74)	(1.78)	(1.83)	(3.23)	-	-	-	-	-	-	-
EGARCH(1,1)	-0.007	-	-11.68*	0.189**	-0.968*	-0.006	-	-	-	-	-	-
	(-1.55)	-	(-41.75)	(2.27)	(-25.43)	(-0.07)	-	-	-	-	-	-
TARCH(1,1)	-0.001	-	0.00015	0.111	0.873*	-	-0.078	-	-	-	-	-
	(-0.25)	-	(0.94)	(1.36)	(9.33)	-	(-0.94)	-	-	-	-	-
PARCH(1,1)	-0.003	-	0.007	0.065	0.916*	-	-	-0.534	0.523	-	-	-
	(-0.87)	-	(0.23)	(0.67)	(7.66)	-	-	(-0.59)	(0.31)	-	-	-
CGARCH(1,1)	-0.00028	-	0.002*	-0.180*	0.223	-	-	-	-	0.927*	0.063	0.494*
	(-0.09)	-	(3.68)	(-2.59)	(0.84)	-	-	-	-	(9.47)	(0.86)	(3.08)

Notes: z-statistics in parentheses; *, ** and *** denote significance at the 1%, 5% and 10% levels, respectively.
The time series of the empirical quarterly spot return itself is obtained from quarterly point spot rates in time.

Table 3.10: Estimates of GARCH class conditional volatility models for the empirical quarterly spot return of USD/GBP itself

same as that in Table 3.8. The variance intercept terms in the GARCH, TARCH and PARCH are not significant, which is slightly similar to that in Table 3.8 where only the variance intercept term in the PARCH model is not significant. In common, the variance intercept terms in GARCH, GARCH-M and TARCH are very small and only the PARCH model has the sum of the ARCH and GARCH coefficients close to unity to capture persistent shocks to conditional variance as in Tables 3.6 and 3.8. Generally, the estimating results for the empirical quarterly USD/GBP spot return itself in Table 3.10 have slightly more similarities to those in Table 3.8 for the empirical monthly USD/GBP spot return itself than those in Table 3.6 for the theoretical quarterly spot return itself.

We report the estimating results of the GARCH class models for the innovation of the quarterly USD/GBP spot return that is of an ARMA(2,3) process[35] in Table 3.11. In Table 3.11, the coefficients on the ARCH, and GARCH terms in the conditional variance equation are statistically significant, except both the non-significant ARCH and GARCH terms in the CGARCH(1,1) model and the non-significant ARCH term in the ARCH model. In particular, all the GARCH coefficient estimates except the one in the CGARCH model are highly statistically significant (at 1%), which is the same as in Table 3.9. The asymmetry terms (η, μ and γ) in the asymmetric EGARCH, TARCH and CGARCH models are significant, which is highly similar to the asymmetric estimates in Table 3.7. The EGARCH, PARCH and TARCH models capture persistent volatility shocks due to the sum of their ARCH and GARCH terms being approximately close to unity. The conditional variance term in the mean equation of the GARCH-M model has a negative sign and is significant at the 5% level, suggesting that higher

[35]We find that the time series of the quarterly USD/GBP spot return constructed from quarterly point spot rates in time is an ARMA(2,3) using Schwarz's (1978) Bayesian information criterion (SBIC) that is recommended by Diebold (2001).

Mean Equation

$$\Delta s_t = c + c_1 \Delta s_{t-1} + c_2 \Delta s_{t-2} + c_3 \tau_{t-1} + c_4 \tau_{t-2} + c_5 \tau_{t-3} + \tau_t \quad \tau_t \backsim N(0, h_t^2)$$

(where $\Delta s_t = c + c_1 \Delta s_{t-1} + c_2 \Delta s_{t-2} + c_3 \tau_{t-1} + c_4 \tau_{t-2} + c_5 \tau_{t-3} + d h_t^2 + \tau_t \quad \tau_t \sim N(0, h_t^2)$ to GARCH(1,1)-M)

Variance Equation

$h_t^2 = \omega + \alpha \tau_{t-1}^2$	ARCH(1)
$h_t^2 = \omega + \alpha \tau_{t-1}^2 + \beta h_{t-1}^2$	GARCH(1,1)
$h_t^2 = \omega + \alpha \tau_{t-1}^2 + \beta h_{t-1}^2$	GARCH(1,1)-M
$\ln(h_t^2) = \omega + \alpha \left[\left| \frac{\tau_{t-1}}{h_{t-1}} \right| - \sqrt{\frac{2}{\pi}} \right] + \beta \ln(h_{t-1}^2) + \eta \frac{\tau_{t-1}}{\lvert h_{t-1} \rvert}$	EGARCH(1,1)
$h_t^2 = \omega + \alpha \tau_{t-1}^2 + \beta h_{t-1}^2 + \mu \tau_{t-1}^2 I_{t-1}$ where $I_{t-1} = 1$ if $\tau_{t-1} < 0$ and 0 otherwise	TARCH(1,1)
$h_t^{\vartheta} = \omega + \alpha (\lvert \tau_{t-1} \rvert - \upsilon \tau_{t-1})^{\vartheta} + \beta h_{t-1}^{\vartheta}$	PARCH(1,1)
$h_t^2 = \omega + \rho (q_{t-1} - \omega) + \varphi (\tau_{t-1}^2 - h_{t-1}^2) + \alpha (\tau_{t-1}^2 - q_{t-1}) + \gamma (\tau_{t-1}^2 - q_{t-1}) D_{t-1} + \beta (h_{t-1}^2 - q_{t-1})$	CGARCH(1,1)

Estimating model	c	c_1	c_2	c_3	c_4	c_5	d	ω	α	β	η	μ	υ	ϑ	ρ	φ	γ
ARMA(2,3)-	-0.001	-0.284*	-0.959*	0.520*	1.023*	0.300*	-	0.002*	0.159	-	-	-	-	-	-	-	-
ARCH(1)	(-0.23)	(-14.44)	(-44.04)	(5.31)	(38.04)	(3.08)	-	(6.23)	(1.23)	-	-	-	-	-	-	-	-
ARMA(2,3)-	-0.002	-0.283*	-0.954*	0.508*	1.023*	0.286*	-	0.00046	0.153***	0.675*	-	-	-	-	-	-	-
GARCH(1,1)	(-0.37)	(-14.84)	(-47.17)	(4.77)	(42.82)	(2.78)	-	(1.47)	(1.89)	(3.90)	-	-	-	-	-	-	-
ARMA(2,3)-	0.137**	-0.277*	-0.957*	0.532*	1.033*	0.313*	-58.345**	0.00064*	0.066*	0.689*	-	-	-	-	-	-	-
GARCH(1,1)-M	(2.17)	(-18.59)	(-57.34)	(6.54)	(51.46)	(3.91)	(-2.38)	(2.58)	(2.86)	(6.69)	-	-	-	-	-	-	-
ARMA(2,3)-	-0.00061	-0.282*	-0.965*	0.525*	1.031*	0.300*	-	-0.769	0.189***	0.899*	0.127***	-	-	-	-	-	-
EGARCH(1,1)	(-0.13)	(-15.66)	(-53.87)	(5.49)	(44.15)	(3.33)	-	(-1.36)	(1.88)	(10.50)	(1.88)	-	-	-	-	-	-
ARMA(2,3)-	-0.00095	-0.285*	-0.965*	0.513*	1.029*	0.275*	-	0.00020	0.188***	0.829*	-	-0.198***	-	-	-	-	-
TARCH(1,1)	(-0.20)	(-16.81)	(-59.85)	(5.64)	(37.95)	(3.29)	-	(1.09)	(1.80)	(7.09)	-	(-1.79)	-	-	-	-	-
ARMA(2,3)-	0.00008	-0.284*	-0.960*	0.539*	1.032*	0.319*	-	0.015	0.109***	0.816*	-	-	-0.634	0.615	-	-	-
PARCH(1,1)	(0.02)	(-15.94)	(-52.30)	(6.30)	(45.82)	(4.04)	-	(0.34)	(1.91)	(6.8)	-	-	(-1.57)	(0.65)	-	-	-
ARMA(2,3)-	-0.00282	-0.213**	-0.746*	0.319*	0.745*	0.278*	-	0.002*	-0.367	0.354	-	-	-	-	0.671**	0.251	0.378**
CGARCH(1,1)	(-0.64)	(-2.00)	(-6.58)	(2.57)	(5.92)	(3.21)	-	(4.45)	(-1.37)	(0.62)	-	-	-	-	(2.32)	(0.91)	(1.99)

Notes: z-statistics in parentheses; *, ** and *** denote significance at the 1%, 5% and 10% levels, respectively.

The time series of the empirical quarterly ARMA(2,3) spot return is obtained from quarterly point spot rates in time.

Table 3.11: Estimates of GARCH class conditional volatility models for the empirical quarterly ARMA(2,3) spot return of USD/GBP

market-wide risk will lead to lower returns. Only in the ARCH, GARCH-M and CGARCH models are the variance intercept terms significant. The variance intercept terms in the GARCH, GARCH-M and TARCH are globally small, as found in Tables 3.6-3.11. Generally, the estimating results for the empirical quarterly USD/GBP ARMA(2,3) spot return in Table 3.11 have slightly more similarities to those in Table 3.7 for the theoretical quarterly ARMA(2,2) spot return than those in Table 3.9 for the empirical monthly USD/GBP MA(1) spot return.

3.7.1.3 Testing significance of estimates

In order to check and make sense of the estimates exposed previously, the final thing we are going to do is to test the significant differences of the empirical estimates from the theoretical estimates. Treating the theoretical estimates as fixed numbers, we take the empirical estimates of the GARCH parameters and their standard errors and apply a t-test to each of the GARCH parameter's significance from that estimated with the theoretical data. The null hypothesis is that the population GARCH parameters (α, β) are the theoretical estimates against the 5% two-sided alternative. We quote the test statistics and report the results in Table 3.12 and Table 3.13.

In Table 3.12, we report the results of testing estimation significance for the spot return itself. Given the t-ratios of the estimated ARCH and GARCH coefficients and 5% two-sided critical values[36], on the one hand, for the empirical monthly spot return itself in the middle part (from the 4th to 9th columns) of Table 3.12, only the CGARCH model has the non-significant ARCH and GARCH parameters while the rest of the models have significant ARCH coefficients and only the EGARCH model has the significant GARCH estimate. The CGARCH model has the best estimating

[36]For the degrees of freedom greater than around 25, the 5% two-sided critical value is approximately ± 2. As a rule of thumb, the null hypothesis would be rejected if the t-statistic exceeds 2 in absolute value.

Spot return itself	Theoretical quarterly spot return		Empirical monthly spot return						Empirical quarterly spot return					
	α_{th}	β_{th}	α_{em}	t_α	*rejection*	β_{em}	t_β	*rejection*	α_{em}	t_α	*rejection*	β_{em}	t_β	*rejection*
ARCH(1)	15.077	-	0.322	-138.81	*yes*	-	-	-	0.153	-119.90	*yes*	-	-	-
	(2.11)	(-)	(0.11)			(-)			(0.12)			(-)		
GARCH(1,1)	0.578	0.420	0.259	-4.03	*yes*	0.609	1.77	*no*	0.073	-9.71	*yes*	0.867	4.06	*yes*
	(0.04)	(0.04)	(0.08)			(0.11)			(0.05)			(0.11)		
GARCH(1,1)-M	0.447	0.551	0.257	-2.36	*yes*	0.607	0.51	*no*	0.153	-3.52	*yes*	0.591	0.21	*no*
	(0.02)	(0.02)	(0.08)			(0.11)			(0.08)			(0.18)		
EGARCH(1,1)	1.017	0.164	0.392	-6.87	*yes*	0.857	10.05	*yes*	0.189	-9.95	*yes*	-0.968	-29.76	*yes*
	(0.52)	(0.26)	(0.09)			(0.07)			(0.08)			(0.04)		
TARCH(1,1)	1.213	0.398	0.215	-12.60	*yes*	0.611	1.98	*no*	0.111	-13.52	*yes*	0.873	5.07	*yes*
	(0.10)	(0.02)	(0.08)			(0.11)			(0.08)			(0.09)		
PARCH(1,1)	0.537	0.605	0.190	-6.13	*yes*	0.732	1.48	*no*	0.065	-4.91	*yes*	0.916	2.60	*yes*
	(0.04)	(0.02)	(0.06)			(0.09)			(0.10)			(0.12)		
CGARCH(1,1)	-0.038	0.514	-0.072	-0.36	*no*	0.701	1.72	*no*	-0.180	-2.05	*yes*	0.223	-1.09	*no*
	(0.04)	(0.23)	(0.10)			(0.11)			(0.07)			(0.27)		

Notes: The t-tests are H_0: $\alpha = \alpha_{th}$, H_1: $\alpha \neq \alpha_{th}$ and H_0: $\beta = \beta_{th}$, H_1: $\beta \neq \beta_{th}$. The test statistics are $t_\alpha = (\alpha_{em} - \alpha_{th})/SE(\alpha_{em})$, $t_\beta = (\beta_{em} - \beta_{th})/SE(\beta_{em})$, respectively. α is the coefficient estimate on the ARCH term. β is the coefficient estimate on the GARCH term. The subscript "$_{th}$" refers to the theoretical data. The subscript "$_{em}$" refers to the empirical data. The time series of the empirical monthly spot return is obtained from monthly averages of daily spot rates in that month. The time series of the empirical quarterly spot return is obtained from quarterly point spot rates in time. Given the t-ratios and critical values (2 for a 5% test), the null hypotheses are rejected. "-" indicates N/A. The standard errors in parentheses are placed below the coefficient estimates.

Table 3.12: Testing significant differences of the empirical estimates from the theoretical estimates for the spot return itself

performance because the estimated values of both the ARCH and GARCH coefficients are indistinguishable statistically from the theoretical estimated values of these two parameters. The EGARCH model has both the statistically distinguishable ARCH and GARCH parameters compared to those of the theoretical estimates. On the other hand, for the empirical quarterly spot return itself in the right hand side (from the 10th to 15th columns) of Table 3.12, all models have the significant ARCH coefficients. The GARCH-M and CGARCH model have the insignificant GARCH parameters while the other models have the significant GARCH estimates. The CGARCH model is one of the best estimating models with less significant differences. To summarize, in Table 3.12, the GARCH class models for the empirical monthly spot return of averaging data have the same or better estimating performance as they do for the empirical quarterly spot return of point data in time. The empirical monthly spot return has less estimation differences of significance to those of the theoretical quarterly spot return than the empirical quarterly spot return.

In Table 3.13, we report the results of testing estimation significance for the spot return of the ARMA process. Given the t-ratios of the estimated ARCH and GARCH coefficients and 5% two-sided critical values, on the one hand, for the empirical monthly MA(1) spot return in the middle part (from the 4th to 9th columns) of Table 3.13, all models except CGARCH have the significant ARCH coefficients while all models except CGARCH and PARCH have the significant GARCH parameters. Only the CGARCH model has the non-significant ARCH and GARCH estimates, which is the same as in Table 3.12 for the empirical monthly spot return itself. The CGARCH model has the best estimating performance because the estimated values of the ARCH and GARCH coefficients are indistinguishable statistically from the theoretical estimated parameter values. The PARCH model is the second best estimating model due to the

Spot return of the ARMA process

	Theoretical quarterly ARMA(2,2) spot return		Empirical monthly MA(1) spot return						Empirical quarterly ARMA(2,3) spot return					
	α_{th}	β_{th}	α_{em}	t_α	*rejection*	β_{em}	t_β	*rejection*	α_{em}	t_α	*rejection*	β_{em}	t_β	*rejection*
ARCH(1)	38.372	-	0.187	-415.28	*yes*	-	-	-	0.159	-293.44	*yes*	-	-	-
	(11.17)	-	(0.09)			(-)			(0.13)			(-)		
GARCH(1,1)	0.463	0.535	0.139	-6.38	*yes*	0.754	2.41	*yes*	0.153	-3.86	*yes*	0.675	0.80	*no*
	(0.02)	(0.02)	(0.05)			(0.09)			(0.08)			(0.17)		
GARCH(1,1)-M	0.711	0.287	0.140	-11.06	*yes*	0.752	5.02	*yes*	0.066	-27.95	*yes*	0.689	3.90	*yes*
	(0.09)	(0.09)	(0.05)			(0.09)			(0.02)			(0.10)		
EGARCH(1,1)	0.948	0.943	0.226	-5.59	*yes*	-0.622	-7.57	*yes*	0.189	-7.55	*yes*	0.899	-0.50	*no*
	(0.02)	(0.01)	(0.13)			(0.21)			(0.10)			(0.09)		
TARCH(1,1)	1.171	0.401	0.130	-19.56	*yes*	0.754	4.00	*yes*	0.188	-9.45	*yes*	0.829	3.67	*yes*
	(0.09)	(0.02)	(0.05)			(0.09)			(0.10)			(0.12)		
PARCH(1,1)	0.526	0.617	0.142	-8.53	*yes*	0.772	1.76	*no*	0.109	-7.29	*yes*	0.816	1.66	*no*
	(0.04)	(0.02)	(0.04)			(0.09)			(0.06)			(0.12)		
CGARCH(1,1)	0.097	0.623	-0.030	-1.68	*no*	-0.345	-1.59	*no*	-0.367	-1.73	*no*	0.354	-0.47	*no*
	(0.02)	(0.02)	(0.08)			(0.61)			(0.27)			(0.57)		

Notes: The t-tests are H_0: $\alpha_{em}=\alpha_{th}$, H_1: $\alpha_{em}\neq\alpha_{th}$ and H_0: $\beta_{em}=\beta_{th}$, H_1: $\beta_{em}\neq\beta_{th}$. The test statistics are $t_\alpha=(\alpha_{em}-\alpha_{th})/SE(\alpha_{em})$, $t_\beta=(\beta_{em}-\beta_{th})/SE(\beta_{em})$, respectively. α is the coefficient estimate on the ARCH term. β is the coefficient estimate on the GARCH term. The subscript "$_{th}$" refers to the theoretical data. The subscript "$_{em}$" refers to the empirical data. The time series of the empirical monthly spot return is obtained from monthly averages of daily spot rates in that month. The time series of the empirical quarterly spot return is obtained from quarterly point spot rates in time. Given the t-ratios and critical values (2 for a 5% test), the null hypotheses are rejected. "-" indicates N/A. The standard errors in parentheses are placed below the coefficient estimates.

Table 3.13: Testing significant differences of the empirical estimates from the theoretical estimates for the ARMA spot return

indistinguishable GARCH coefficient. On the other hand, for the empirical quarterly ARMA(2,3) spot return in the right hand side (from the 10th to 15th columns) of Table 3.13, all models except CGARCH have the significant ARCH coefficients while the GARCH, CGARCH, EGARCH and PARCH models have the non-significant GARCH parameters. Again, the CGARCH model is the only model that has the statistically indistinguishable ARCH and GARCH parameters from those theoretical estimated values. Generally, in Table 3.13, the GARCH class models for the empirical quarterly ARMA(2,3) spot return of point data in time have the same or better[37] estimating performance as they do for the empirical monthly MA(1) spot return of averaging data. The empirical quarterly ARMA(2,3) spot return has less estimation differences of significance to those of the theoretical quarterly ARMA(2,2) spot return than it does to the empirical monthly MA(1) spot return. The CGARCH, PARCH and EGARCH models are the top three best estimating models.

Overall, taking account of the theoretical and empirical estimates of the significance of coefficients (especially those on the GARCH terms), asymmetric effects, additional ARCH, and persistent volatility shocks and the results of testing estimated significant differences in Tables 3.6-3.13, it is found that the best fit estimating models for both the theoretical and empirical data are the asymmetric CGARCH, EGARCH, and PARCH conditional volatility models, which is consistent with what we find about forecast in Chapter 2[38]. The theoretical model can produce the same kind of ARCH estimates as we see in the real-world data.

[37]For the empirical quarterly spot return, the ARCH, GARCH-M, CGARCH, PARCH and TARCH models have the same estimating performances and the GARCH and EGARCH models have the better performances of the ARCH and GARCH parameters compared to theirs for the empirical monthly spot return.

[38]In Chapter 2, although we conclude that no single model dominates for forecasts, the CGARCH, EGARCH and PARCH models have locally the best forecasting performances compared to those of other symmetric and asymmetric GARCH class models.

3.7.2 Impulse response function

We establish the dynamic form and analyze the dynamic properties of conditional volatility estimated by examining the impulse response function (IRF). We are interested in knowing 'how does a unit innovation to conditional volatility affect it, now and in future?' An impulse response function measures the effect of a transitory shock to current volatility h_0^2 on future volatilities h_t^2. To achieve this, we read off the coefficients in the moving average representation of the process. For example, we consider the GARCH(1,1) model and its conditional variance process is defined by

$$h_t^2 = \omega + \alpha\tau_{t-1}^2 + \beta h_{t-1}^2 \qquad \tau_t \backsim N(0, h_t^2)$$

Subtracting 1 from each of the time subscripts, an infinite number of successive substitutions of the conditional variance would yield

$$h_t^2 = \omega(1 + \beta + \beta^2 + ...) + \alpha\tau_{t-1}^2\left(1 + \beta L + \beta^2 L^2 + ...\right) + \beta^{\infty} h_0^2$$

The expression on the RHS is simply a constant, and as the number of observations tends to infinity, β^{∞} will tend to zero. Hence, equivalently, the volatility equation can be written as

$$h_t^2 = \varpi + Z_1\tau_{t-1}^2 + Z_2\tau_{t-2}^2 + Z_3\tau_{t-3}^2 + ...Z_t\tau_0^2$$

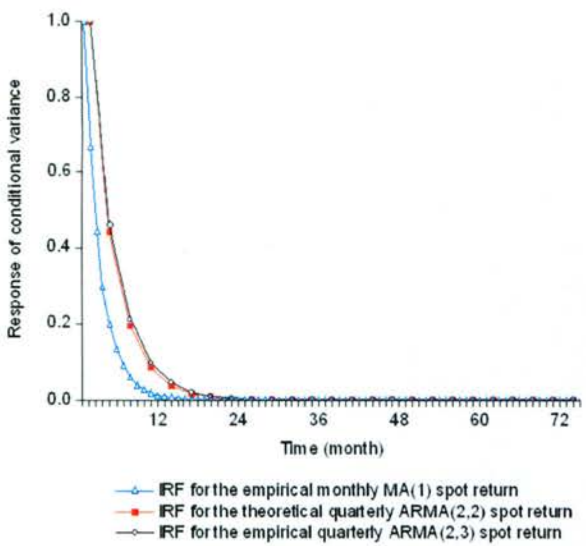
where ϖ is the constant term, $Z_1 = a$, $Z_2 = \alpha\beta$, $Z_3 = \alpha\beta^2$, $Z_4 = \alpha\beta^3$, ..., $Z_t = \alpha\beta^{(t-1)}$. The full set of impulse-response coefficients, $\{Z_1, Z_2 ... Z_t\}$, tracks the complete dynamic response of h_t^2 to the shock. In other words, the autoregressive model is written as a moving average.

Using the IRF we can convert the conditional variance (h_t^2) in the variance equation of our best fit estimating CGARCH(1,1), EGARCH(1,1) and PARCH(1,1) models for the theoretical (simulated) ARMA(2,2) spot return (Δs_t) into their infinite moving average forms, respectively. The IRF is then simply the graph of Z_i against $i = 1, 2, 3....$ We repeat this process for the best estimating models for the conditional variance of the empirical monthly USD/GBP spot return that is an MA(1) and for the conditional variance of the empirical quarterly USD/GBP spot return that is an ARMA(2,3). We start the IRFs for the theoretical and empirical data at the same point. Three IRFs start from the same initial response (e.g. one unit shock) and then any differences between the dynamic patterns of the three can be seen clearly. We draw and compare the impulse response functions (IRFs) for the conditional variance (h_t^2) of the three ARCH processes we estimated. It is noted that we compress the scale so the monthly IRFs from the empirical data are synchronized with the quarterly IRFs from the theoretical data. For the latter, we need the IRFs using the theoretical model's parameters in the baseline as in Table 3.1 to get the theoretical time series of the ARMA(2,2) spot return we simulated.

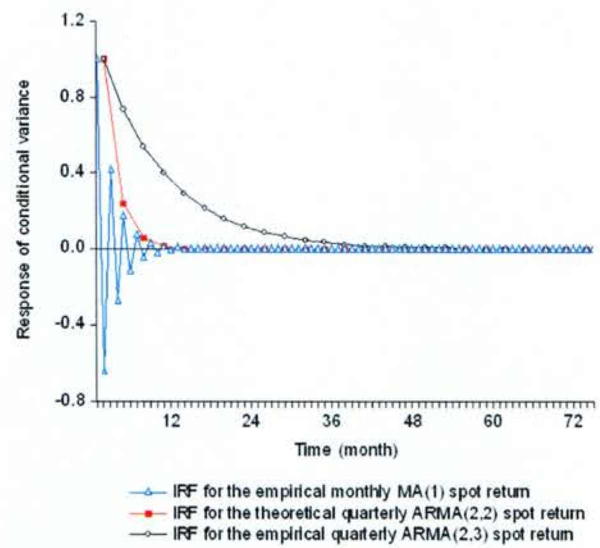
In Figure 3.1, we display the impulse responses of conditional variance of the best fit estimating models at time t to one unit shock in variance at time 0 for the theoretical and the empirical ARMA spot returns, respectively. First, we consider the sign of the responses. For the empirical monthly IRFs from the the empirical monthly MA(1) spot return (blue lines), the line graphs on the left and right hand sides (LHS and RHS) show that a shock to the conditional variance in the CGARCH and PARCH models respectively always has a positive impact on the future conditional volatility since the impulse responses are positive. The line graph in the middle position shows that a shock to the conditional variance in the EGARCH model has a positive (negative) impact at the odd (even) time points on the future conditional volatility since the impulse response is positive (negative) at that time. For the theoretical quarterly IRFs from the theoretical quarterly ARMA(2,2) spot return (red lines), for all three cases,

Innovations in conditional variance of white noise errors

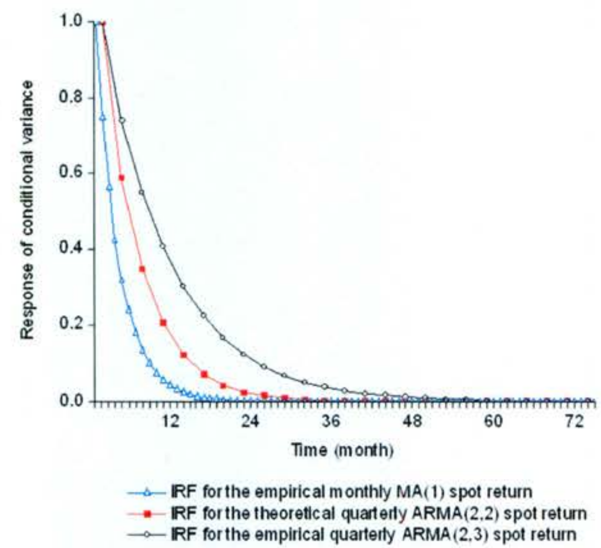
Response of Conditional Variance in CGARCH(1,1) to One Unit Innovation



Response of Conditional Variance in EGARCH(1,1) to One Unit Innovation



Response of Conditional Variance in PARCH(1,1) to One Unit Innovation



Notes: The time series of the empirical monthly MA(1) spot return is obtained from monthly averages of daily spot rates in that month; the time series of the empirical quarterly ARMA(2,3) spot return is obtained from quarterly point spot rates in time.

Figure 3.1: Impulse response functions for one unit innovation in conditional variance from the empirical and the theoretical ARMA spot returns

the conditional volatility always has a positive response to shock at all time points until the effect of the shock dies out. The empirical quarterly IRFs from the empirical quarterly ARMA(2,3) spot return (black lines) have positive signs for all three cases, which is the same as the theoretical quarterly IRFs have. The empirical quarterly IRFs seem to be having more similarities to the theoretical quarterly IRFs than the empirical monthly IRFs in terms of the sign of the responses.

Second, we deliberate on the magnitude of the responses and the time periods the effect of the shock takes to die out. In Figure 3.1, the monthly impulse responses are small while the quarterly impulse responses are big. For the empirical monthly IRFs from the empirical monthly MA(1) spot return (blue lines), for all three cases, the effect of the shock takes approximately $1.5 - 2$ year to die out. In details, the effect of the shock to the conditional variance in the CGARCH, EGARCH and PARCH models dies down toward zero (≈ 0.001) after taking approximately 17, 15, and 23 months, respectively. In contrast, for the theoretical quarterly IRFs from the theoretical quarterly ARMA(2,2) spot return (red lines), the effect of the shock to the conditional variance in the CGARCH, EGARCH and PARCH models takes 24, 15, and 39 months respectively to die out, while, for the empirical quarterly IRFs from the empirical quarterly ARMA(2,3) spot return (black lines), the effect of the shock to the conditional variance in the CGARCH, EGARCH and PARCH models takes 27, 66, and 66 months respectively to die down toward zero (≈ 0.001). For the CGARCH model, the empirical quarterly IRF has the highly similar magnitudes of responses and time periods for dying out as the theoretical quarterly IRF has so that both IRFs look almost same (overlapping). For the EGARCH and PARCH models, the empirical monthly IRFs have more similarities to the theoretical quarterly IRFs than the empirical quarterly IRFs. The quarterly IRFs take longer than the monthly IRFs to die down, whereas the empirical quarterly IRFs with the biggest magnitudes take longest time. All IRFs from the empirical and the theoretical ARMA spot returns in Figure 3.1 show a clear trend of dying out, where the effect of the shock to the conditional variance in the CGARCH model takes the shortest time while the one to the PARCH model takes the longest time, in short, $CGARCH < EGARCH < PARCH$.

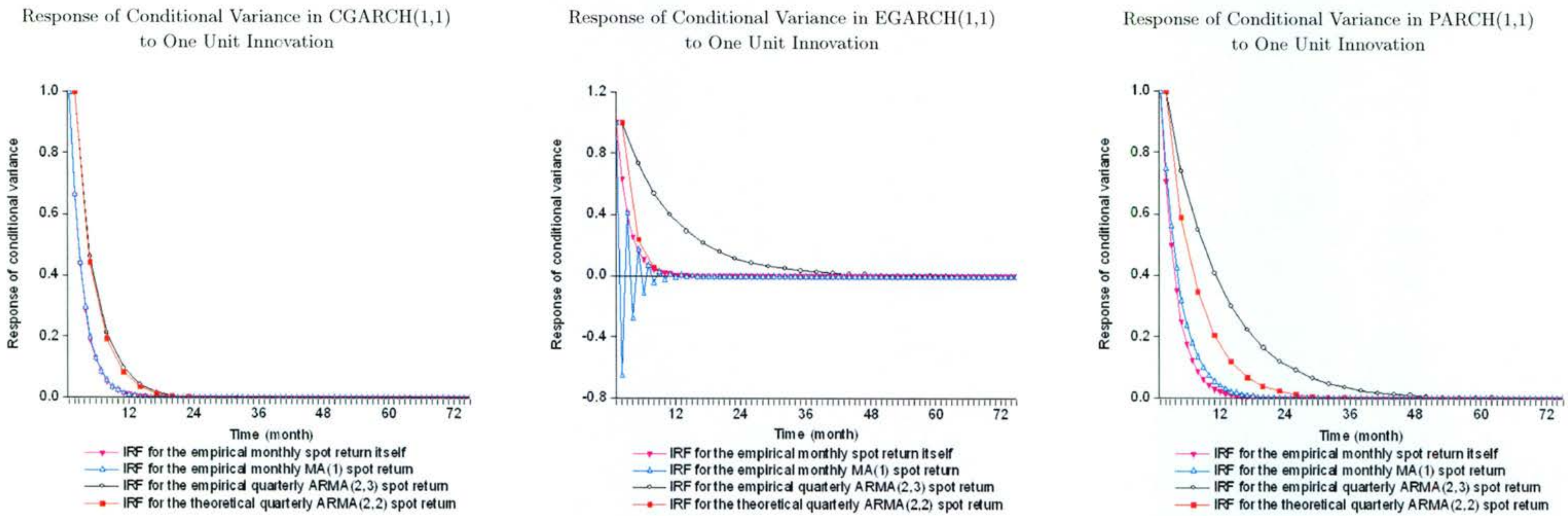
Third, we consider the dynamic features of the responses. For the empirical monthly IRFs from the empirical monthly MA(1) spot return (blue lines), the impulse responses of the conditional variance in the CGARCH and PARCH models on the LHS and RHS respectively in Figure 3.1 monotonically decrease up to 72 months until the effect of the shock dies out. Specifically, the IRF to the CGARCH model tends to decline more intensely than the IRF to the PARCH model does within the first 12 months particularly on the time interval [4, 12], while the impulse responses of the conditional variance in the EGARCH model in the middle part fluctuate around zero with a gradually (from strong to weak) falling trend as the effect of the shock dies out. All theoretical quarterly IRFs from the theoretical quarterly ARMA(2,2) spot return (red lines) decay in a monotonic decreasing fashion. It is found in common that both the empirical monthly and the theoretical quarterly IRFs to the CGARCH and PARCH models drop slowly with a relatively smooth trend, while both the empirical monthly and the theoretical quarterly IRFs to the EGARCH model descends quickly with a comparably steep trend. Unlike the dynamic movements of the empirical monthly and the theoretical quarterly IRFs, the empirical quarterly IRFs from the empirical quarterly ARMA(2,3) spot return (black lines) have the more intense decline with the CGARCH model than with the EGARCH and PARCH models, in which the empirical quarterly IRFs are quite smooth with a generally falling trend. In Figure 3.1, the shock has a diminishing impact on future conditional volatility, and all of the empirical and theoretical IRFs show a (approximately) monotonic decreasing fashion until they finally die out.

It is emphasized that we are interested in the IRF's dynamics not the size of the typical shock (e.g. variance of shocks) "hitting" the variances. The empirical and theoretical IRFs would differ anyway because we compare monthly average data against quarterly point data in time. For the empirical and theoretical data, the dynamic patterns of the CGARCH and PARCH models respectively always look closer on the IRFs than those of the EGARCH model, which, as we expect, is consistent with the results of testing significance of the data estimates disclosed previously where the CGARCH and PARCH parameters of the empirical data estimates are not significantly different from the ones of theoretical data.

As reported previously, the results of the ARCH estimates for the empirical monthly spot return itself (without an ARMA process) have more similarities to those for the innovation in the theoretical quarterly spot return of an ARMA(2,2) process. We plot the IRFs for our best fit estimating (CGARCH, EGARCH and PARCH) models for the conditional variance of the empirical monthly spot return itself due to its superior ARCH estimates. In Figure 3.2, we compare these IRFs with the IRFs implied by the empirical and theoretical ARCH processes from the ARMA spot returns as in Figure 3.1.

In Figure 3.2, for the empirical monthly IRFs from the empirical monthly spot return itself (mauve lines), the impulse responses are positive and smaller than both the empirical and theoretical quarterly impulse responses (red and black lines), while they are not greater ($\leqq$) than the absolute values of the IRFs' magnitudes from the empirical monthly MA(1) spot return (blue lines) at all time points for all model cases. The effect of the shock to the conditional variances in the CGARCH, EGARCH and PARCH models takes approximately 16, 15, and 19 months respectively to die out.

Innovations in conditional variance of white noise errors



Notes: The time series of the empirical monthly spot return itself is obtained from monthly averages of daily spot rates in that month; the time series of the empirical monthly MA(1) spot return is obtained from monthly averages of daily spot rates in that month; the time series of the empirical quarterly ARMA(2,3) spot return is obtained from quarterly point spot rates in time.

Figure 3.2: Impulse response functions for one unit innovation in conditional variance from the empirical and the theoretical spot returns

The shock has a diminishing impact on future conditional volatility, and the impulse responses of the conditional volatility to shock tend to have a monotonic decreasing fashion, which is consistent with the findings in Figure 3.1. It is seen clearly that, for all cases, the IRFs in mauve monotonically decrease at a similar steep slope as the IRFs from the theoretical ARMA(2,2) spot return do, which is better than the IRFs in blue (from the empirical monthly MA(1) spot return) though the IRFs in blue are superior to those in black (from the empirical quarterly ARMA(2,3) spot return). It suggests that the dynamic forms of conditional heteroscedasticity for the empirical monthly spot return itself without an ARMA process have more similarities, meaning it looks more consistent with those for the innovation in the theoretical quarterly ARMA(2,2) spot return. This, we think, could be another justification for empirical researchers considering the FOREX spot return itself not its innovation.

We also look at the dynamic movements of the impulse responses to the shock for other types of empirical data in the quarterly frequency where two time series of the empirical quarterly spot return itself are constructed from the empirical quarterly averages of daily USD/GBP spot rates in the quarter and from the empirical quarterly USD/GBP point spot rates in time, respectively; and one time series of the empirical quarterly ARMA(2,1) spot return is constructed from the quarterly averages of daily USD/GBP spot rates in that quarter. We find, but do not report, that, as the consistent result, all IRFs tend to have the monotonic decreasing fashion and the empirical monthly spot return itself constructed from the averaging data has the most similar ARCH dynamics to those of the theoretical quarterly data.

Overall, as shown in both Figures 3.1 and 3.2, the empirical monthly IRFs "look the same" as the theoretical quarterly IRFs for an approximately monotonic decreasing fashion as time goes by. Shock has a weakening impact on future conditional volatility until it dies out. The IRFs for variance for the theoretical and empirical data do have similar dynamics, although they have different orders of magnitude, in which the theoretical IRFs are bigger and take longer time to die out. As we mentioned before, earlier on in the chapter, it is dynamics we are interested in not the size of the typical shock "hitting" the variances. The magnitudes would differ anyway because the monthly average data is compared against the quarterly point data in time. In detail, the empirical data uses the average exchange rates whereas the theoretical data is a one-point-in-time observation. Variances of averages are always lower because averaging a series smooths it out. It may be this that is giving the bigger responses to the theoretical model. At the same time, another possible explanation might be the "imperfect" market in reality. In the FOREX markets, many endemic and exotic factors coexist simultaneously. Some of them counteract each other, or some earlier shocks (e.g. old news) are of little or no effect on future conditional volatility, while these factors are not even considered in the pure theoretical framework. For example, investors usually pay more attention to the recent release or announcement of financial news and the market could be influenced mainly by the late central bank intervention. The theoretical model can generate the same kind of dynamic features to ARCH as we see in the actual data.

3.8 Conclusion

In the chapter, we attempt to give a theoretical underpinning to the well established empirical stylized fact that asset returns in general and the spot FOREX returns in particular, display predictable volatility characteristics. After investigating Moore and Roche's habit version of Lucas to conditional volatility, we find that the Lucas two-country, two-good, two-money economy model with habit can generate realistic conditional volatility in the spot FOREX return. Specifically, we research the Lucas two-country monetary model with habit in Moore and Roche (2006) and find the model's implied property of the ARMA(2,2) spot return. We numerically solve the model and test that the theoretical ARMA(2,2) spot return and its innovation in the spot return process are definitely conditionally heteroscedastic. We estimate the best fit GARCH models for the theoretical and empirical data and then establish the dynamic form for the conditional heteroscedasticity we estimate from the best fit GARCH models. Using the impulse response functions (IRFs) we show that the baseline theoretical data has "ARCH" properties in the quarterly frequency that match well the "ARCH" properties of the empirical monthly estimations. The IRFs for the ARCH processes we estimate "look the same" with similar impulse responses to one unit shock in conditional variance of white noise errors. The impulse responses of conditional variance to shock tend to monotonically decrease until the effect of the shock dies out. On the other hand, concerning the highly consistent performance of the ARCH estimates and dynamic features between the empirical monthly spot return itself and the innovation in the theoretical quarterly ARMA(2,2) spot return, we explain why empirical researchers tend to consider the spot FOREX return itself rather than its innovation, as we did in Chapter 2. The Lucas two-country monetary model with habit is capable of producing the same kind of ARCH features as we see in the real data.

As one of the theoretical asset pricing models, the habit persistence model is able not only to explain persistent volatility on asset returns that is unconditional but also to generate volatility clustering in FOREX returns, even its asymmetric properties. Campbell and Cochrane (1999) were the earlier to explain the dynamic behaviour of asset prices using a consumption-based asset pricing model with an external habit. Moore and Roche (2006) extended this theoretical model to a two-country monetary economy to solve many FOREX puzzles simultaneously, including mimicking the FOREX unconditional volatility. McQueen and Vorkink (2004) is an important paper regarding the application of the theoretical asset pricing model to the issue of volatility clustering. McQueen and Vorkink (2004) developed a preference-based equilibrium asset pricing model, which derives utility from both consumption and financial wealth to endogenously explain conditional volatility in US stock data. In McQueen and Vorkink (2004), a unique mental scorecard that records wealth changes and affects investors' level of risk aversion induces wealth-varying sensitivity to news causing subsequent stock volatility.

In the chapter, we use the theoretical Lucas two-economy representative-agent model in Moore and Roche (2006), which combines the external habit in Campbell and Cochrane (1999) with a monetary framework to explain conditional volatility in spot FOREX returns. According to the capacity of the theoretical asset pricing models in Moore and Roche (2006) and McQueen and Vorkink (2004) to explain volatility clustering, we summarize their main features as follows: 1) Moore and Roche (2006) derive utility from surplus consumption while McQueen and Vorkink (2004) derive utility from consumption and financial wealth changes; 2) Moore and Roche (2006) use an external scorecard of surplus consumption ratio while McQueen and Vorkink (2004) use an internal scorecard of prior investment performance; 3) Moore and Roche (2006) numerically solved the model using the quarterly calibrated parameters while McQueen and Vorkink (2004) numerically solved the model using the monthly calibrated parameters; 4) Moore and Roche (2006) mimic unconditional volatility in FOREX changes while McQueen and Vorkink (2004) explain volatility clustering in asset returns; 5) for both cases, we are moving away from a utility function (U) that can be written as the sum of discounted one-period utility (u) in "current" consumption.

In our opinion, both utility specifications in Moore and Roche (2006) and McQueen and Vorkink (2004) are at the heart of generating ARCH effects, which is overall consistent with what Cochrane (2001) suggests, "risk aversion depends on the level of consumption or wealth relative to some trend or the recent past". In other words, either surplus consumption utility in Moore and Roche (2006) or prior investment utility in McQueen and Vorkink (2004) could be the reason behind the volatility clustering found in empirical facts. However, we think that the utility function in McQueen and Vorkink (2004) is strange. It includes wealth (changes) directly in utility implying that consumers care about wealth directly. But in economics we always think of wealth as an instrument that leads to utility via its ability to buy consumption, rather than the object itself. To take an extreme case, would we be happy including a utility that was a function of Treasury Bill holdings? Of course there is a precedent – money has been included in the utility function in macroeconomics. But this is more of a device rather than something couched in a belief that money itself (rather than consumption) gives you utility. By contrast only consumption appears in a habit utility. It is true however that the habit term collects together past consumptions perhaps in a way that wealth could collect together future consumptions. But at least it is directly in terms of consumption whereas with wealth we would need to convert to consumption via e.g. current and future interest rates and maybe current and future (consumption) price levels. Hence, it is unreasonable to assume that people care about their changes in wealth separately in addition to the consumption stream that is affected by what the changes in wealth bring.

Finally, the habit persistence model is an industry standard now in macroeconomics and finance. Special attention is given to the role of habit persistence in explaining the equity premium puzzle, additional asset-pricing puzzles such as the risk-free-rate puzzle and the forecastability of excess returns (see, for example, Campbell and Cochrane, 1999), many exchange rate puzzles such as disconnect, forward bias, and Meese-Rogoff puzzles including mimicking unconditional volatilities of exchange rates and spot returns etc. (see, for example, Moore and Roche, 2006), observed business-cycle fluctuations and inflation dynamics, and in generating a theory of counter-cyclical markups of prices over marginal costs. The chapter gives a study of the ability of habit persistence to account for conditional volatility in spot FOREX returns.

Part III

ARCH Evaluation

Chapter 4

Forecasting volatility: Optimal forecast error criterion for utility-based loss functions

4.1 Introduction

Volatility forecasting is an important task in financial markets, and it has held the attention of academics and practitioners over the last two decades.

Given the vast number of models available, economic agents must decide which volatility forecasts to use as well as the evaluation criterion upon which to base that decision. It follows that the correct way to evaluate forecasts is to consider and compare the realized values of different decisions made from using alternative sets of forecasts. In contrast to the efforts made in the construction of volatility models and forecasts, little attention has been paid to forecast evaluation in the volatility forecasting literature (Poon and Granger, 2005). The volatility forecast is often compared to a measure of realized volatility.

The literature that compares the relative performance of competing volatility models is either centered around a statistical loss function or an economic loss function (see González-Rivera, Lee and Mishra, 2004). Xekalaki and Degiannakis (2005) define loss function as "measuring either the distance between actual and predicted values or the benefit from the use of these forecasts". The forecaster's objective is to minimize the expected loss. Forecast evaluation, when considered at all, is in terms of the statistical accuracy measures of point forecasts.

The statistical approach requires no economic assumptions and is thus more practical (see Christodoulakis and Satchell, 1998). Statistical loss functions representing the penalty for the difference between the outcome and its prediction are some simple function of the forecast errors. Being the best-known criteria, the statistical loss functions have well-known statistical properties. The preferred statistical loss functions which are motivated by statistical convenience are based on moments of forecast errors. Econometricians and statisticians use mean absolute error (MAE), mean squared error (MSE), mean absolute percent error ($MAPE$), and forecast error variance etc. For example, MSE is just squared-bias plus variance and minimizing this is known to be statistically consistent in many applications. Numerous studies have used MSE to evaluate the performance of volatility models (see, for example, Akgiray, 1989; Pagan and Schwert, 1990; and Bollerslev and Ghysels, 1996). One forecasting method is more accurate than another if its average statistical loss is less. The best model minimizes a function of the forecast errors. There are many econometric forecasting studies that evaluate the model's success using statistical loss functions. Poon and Granger (2003) review a detailed record of volatility forecasting loss functions and relative references. It is found that models classed as accurate due to small statistical loss are not useful in practical situations, and many of the apparent differences in accuracy across methods may not be statistically significant, as there are often a small number of independent out-of-sample forecast errors. Gerlow, Irwin and Liu (1993) show that the accuracy of forecasts according to traditional statistical criteria may give little guide to the potential profitability of employing those forecasts in a market trading strategy. Models that perform poorly on statistical grounds may still yield a profit if used for trading, and vice versa.

Some authors have evaluated the performance of volatility models with criteria based on economic loss functions. Economic loss functions measure the impact of forecasting mistakes upon financial decisions (see Taylor, 2005). As suggested by Bollerslev et al. (1994), economic loss functions that explicitly incorporate the costs faced by volatility forecast users provide the most meaningful forecast evaluations. The utility-based economic criterion, which depends on the use of the estimated volatility, is an appropriate measure of investment performance by risk averse utility maximizers, based on the assumption that an estimator or model of a conditional variance is preferred if, on average, over many time periods, it leads to higher expected utility. For example, West, Edison, and Cho (1993) considered the problem of portfolio allocation based on models that maximize the utility function of the investor. Engle, Kane, and Noh (1997) and Noh, Engle, and Kane (1994) considered different volatility forecasts to maximize the trading profits in buying/selling options. Engle and Mezrich (1997) proposed a quadratic utility function to evaluate different volatility models used in Value-at-Risk calculations. Lopez (2001) considered probability scoring rules that were tailored to a forecast user's decision problem and confirmed that the choice of loss function directly affected the forecast evaluation of different models. Brooks and Persand (2003) evaluated volatility forecasting in a financial risk management setting in terms of value-at-risk (VaR). González-Rivera, Lee and Mishra (2004) compared out-of-sample predictive ability by maximizing utility in stock returns, while they found that the models' relative performance varies with users' evaluation criteria (including a quadratic utility-based loss function). Using these utility-based economic loss functions may well lead to high profits in trading but no econometrician or statistician would recognize it or be able to analyze its statistical properties. The common feature of these papers is that none of them has studied the statistical properties of utility-based economic loss functions so as to contribute to decision making under uncertainty. Poon and Granger (2003) point out that economic loss functions require some assumptions about the costs and benefits of the results and the shape and property of the utility function which are not available in practice. This shows difficulty of recognizing the statistical properties of expected utility or profit maximization evaluation criteria.

Conditional heterscedasticity in financial data is acknowledged in the fact that the financial asset return has thicker tails than a normally distributed variable as defined in ARCH models. One of the primary objectives in ARCH models is to obtain out-of-sample forecasts of the conditional second moments[1] of a process as well as to gain further insight on the uncertainty of forecasts of its conditional mean. The underlying principle in the ARCH class of models is that they explain the random variation in conditional variance and thus reserve this tail thickening (Christodoulakis and Satchell, 1998). A fundamental question is what criteria should one use to judge the superiority of a volatility forecast.

Following the seminar papers of Engle (1982) and Bollerslev (1986) the ARCH and GARCH models are now widely used in economics and finance. Although ARCH models appear to provide a very good in-sample fit, the numerous studies (see, for example Chapter 2; Tse, 1991; Figlewski, 1994; Xu and Taylor, 1995) have led to the

[1] There are four moments of a distribution: the first moment is the mean; the second moment is the variance; the third moment is the skewness; and the fourth moment is kurtosis.

perception that ARCH-type models provide poor out-of-sample forecasts of volatility. Christodoulakis and Satchell (1998) claim that this result relies on the use of traditional forecast evaluation criteria concerning the accuracy and the unbiasedness of forecasts. They show that the inherent noise in the approximation of the actual and unobservable volatility by the squared return results in a misleading forecast evaluation. In detail, the misestimation of traditional forecast performance measures is likely to be worsened by non-normality known to be present in financial data, which is in conflict with the GARCH model assumptions and normal errors. Hence, the approximation of the true volatility by the squared return introduces a substantial noise, which effectively inflates the estimated forecast error statistics and removes any explanatory power of ARCH volatility forecasts with respect to the true volatility. Andersen and Bollerslev (1998) and Christodoulakis and Satchell (1998) find that the use of squared return as volatility proxy lead to relatively high forecast error statistics as well as low R-square and undermine the inference regarding forecast accuracy. Poon and Granger (2005) explain that the standard error of the error statistics will be large because of the difficulty in estimating the fourth moment for thick tails. As the most comprehensive study on ARCH forecast evaluation, Christodoulakis and Satchell's (1998) concludes that non-linear and utility-based evaluation criteria can be more suitable and reliable than the traditional statistical methods of ARCH volatility forecast evaluation using squared returns as a proxy for unobserved volatility. For the relevant literature review, see Christodoulakis and Satchell (1998) for the details.

Our motivation stems from the fact of poor-out-of-sample ARCH forecasting performance when judged on the basis of traditional forecast accuracy criteria versus its good performance when more advanced procedures such as utility-based criteria are employed, whilst these economic evaluation criteria would not be practical if none of their statistical properties is available in econometric or statistic analyses. González-Rivera, Lee and Mishra (2004) find which ARCH model is implied by maximizing utility. Another interesting question is what statistical criterion (based on forecasting errors) does the utility maximization correspond to?

The global aim of the chapter is to find an optimal forecast error criterion which is an approximation to the utility function of ARCH one-step-ahead forecast error in the sense that the expected utility should be a declining function of forecast error. Both the quadratic and exponential utility functions considered are asymmetric and depend on risk aversion parameters and variances etc., where underestimates of conditional volatility lead to lower expected utility than equivalent overestimates. However, it is found that both utility cases are in terms solely of recognizable items like MAE, MSE etc. when we use a regression of expected utility on forecast error for data point where the utility functions are in terms of call option price forecast errors which depend on volatility forecast errors. The averaged form of the regression is the approximate optimal forecast error criterion for the particular investor facing the particular investment decision. It is emphasized that the coefficients in the regression depend on the parameters in the economic problem the investor faces including the risk aversion parameter and the level of conditional variance. The idea is to do this procedure for different levels of risk aversion and see how the regression coefficients change when the risk aversion parameter changes. The functional form of the optimal forecast error criterion is numerically established with the mapping from errors to wealth under the trading rule. These errors enter in a nonlinear and analytically intractable way into the economic loss functions. In the chapter, we try and figure out the statistical model selection criteria based on functions of forecast errors that would roughly equate with maximizing utility. We do this via a mixed analytical and numerical approach. The optimal forecast error criterion lead us to select a volatility model which is very close to that we get from using economic loss function with similar loss.

In the real world, forecasts of volatility are made for a purpose and the relevant purpose in economics is to help decision makers improve their decisions in arbitrage, hedging, risk and financial management under uncertainty. For example, options traders require asset volatilities to price options, and central banks or international investors forecast exchange rates to make financing/investment decisions. In the academic literature there are frequent mentions of this viewpoint but few attempts to carry it forward into a practical example. Standard forecasting textbooks do not discuss the decision-making aspects. See, for example, Box and Jenkins (1970), Granger and Newbold (1986), and Clements and Hendry (1998, 1999). An early discussion of the usefulness of the decision approach is in a book by Theil (1960) whose Sections 8.4 and 8.5 are similar in spirit to our discussion, although quite different in technique. Another important early reference is White (1966) who considers decision theory for forecast evaluation in the dynamic stochastic programming literature. Granger and Pesaran (2000) review some of the techniques developed for forecasts in decision theory for evaluation and consider their usefulness in economics. There are, of course, many papers and books on decision theory but we have found very few specific references to forecast evaluation.

Our main contribution is to propose an optimality criterion of forecast errors for utility maximization under asymmetric loss and, based on it, provides a simple rule to make economic and financial decisions under uncertainty. Rather than imposing a single statistical criterion for all purposes, this framework permits the forecast user to tailor the criterion to their actual decision problem. Also, the properties of loss optimal utility are studied and difficulties arising in practical situations are dealt with.

The remainder of the chapter is organized as follows. The next section presents the data and forecasting models employed in the study, while the utility functions are described briefly in the third section. Optimal forecast error criterion are outlined and discussed in the fourth section with results given in the fifth section. The final section summarizes the chapter, and offers some concluding remarks.

4.2 Data and forecast models

4.2.1 Option pricing

A European call option gives the holder the right, not obligation, to buy the underlying asset at the strike price on the option maturity date agreed in the contract. The famous formula of Black and Scholes (1973) for pricing European call options[2] is derived from the assumed geometric Brownian dynamics for the asset prices and several further assumptions. These include constant interest rates and dividend yields, short selling opportunities, no transaction costs, no taxes, and continuous trading of the asset and the option. The key insight that leads to the formula is the assumption that no one can make arbitrage profits by owning a portfolio that contains variable quantities of the asset and the option. Under these assumptions, the fair price of a call option is given by a function of six parameters: the asset price S, the time until expiry T, the exercise price X, the risk-free interest rate r_f, the dividend yield q_t, and the volatility σ. The general Black-Scholes call formula[3] at time $t+1$ given the information available at time t is

$$C_{BS,t} = S_t e^{-q_t T} N(d_1) - X e^{-r_{f,t} T} N(d_2) \tag{4.1}$$

[2]Most pricing formulae for European options develops basing upon a risk-neutral valuation methodology. Rational option price were first derived by Black and Scholes (1973) and Merton (1973), who assumed asset prices follow a geometric Brownian motion process. There are two main types of options: calls and puts. The general Black-Scholes put formula is $P_{t+1,t} = Xe^{-r_t T}N(-d_2) - S_t e^{-q_t T}N(-d_1)$.

[3]The original formulae in Black and Scholes (1973) assume there are no dividends and hence omit q. These formulae are given by replacing q by zero in the equations that follow.

where

$$d_1 = \frac{\log(S_t/X) + (r_{f,t} - q_t + \frac{1}{2}\sigma^2_{\tau,t})T}{\sigma_{\tau,t}\sqrt{T}}$$
$$d_2 = d_1 - \sigma_{\tau,t}\sqrt{T}$$

where $C_{BS,t}$ is the one-period ahead predicted price of the call option at time t that expires in T periods; S_t is the price of the underlying stock at time t; τ is the time at which the option expires; T is the option remaining time to maturity, $T = \tau - t$; $r_{f,t}$ is the risk-free interest rate at time t; q_t is the dividend yield on the underlying stock at time t; X is the strike stock price; $N(d)$ is the cumulative normal distribution function; and $\sigma_{\tau,t}$ is the volatility of the underlying stock price during the life of the option, which, in the chapter, is the average volatility from time $t+1$ until the maturity date τ given the information at time t. See Black and Scholes (1973), Merton (1973) and Hull (2003) for the further details.

The pricing of options is a cornerstone of financial literature. The Black-Scholes option pricing model is a very important and useful model in estimating the fair value of an option. The approach can be used to price any security whose payoffs depend on the prices of other securities. The main idea is to create a costless self-financing portfolio strategy, whereby long positions are completely financed by short positions, which can replicate the payoff of the derivative. Under the no-arbitrage condition, the dynamic strategy reduces to a partial different equation subject to a set of boundary conditions, which are determined by the specific terms of the derivative security.

Given that S, X, r_f, and T to the Black-Scholes pricing formulae are observable except the volatility parameter σ, once the market has produced a price for the option, a backward induction technique can be used to derive σ. This revealed value is a natural forecast of future volatility. We call this implied volatility (IV). These are the volatilities implied by option prices observed in the market. In practice, traders usually work with implied volatilities. Since the reference period is from t to τ in the future, option implied volatility is often interpreted as a market's expectation of volatility over the option's maturity. The implied volatility for a European call option, traded at the market price C_{market}, is the number $\sigma_{implied}$ that solves the equation

$$C_{market} = C_{BS}(S, T, X, r_f, q, \sigma_{implied}) \tag{4.2}$$

where the solution is unique and C_{BS} increases when σ increases, keeping all other inputs fixed. Implied volatility covaries with realized volatility (Latane and Rendleman, 1976 and Chiras and Manaster, 1978).

Implied volatilities are used to monitor the market's opinion about the volatility of a particular stock. Traders like to calculate implied volatilities from actively traded options on a certain asset and interpolate between them to calculate the appropriate volatility for pricing a less actively traded option on the same stock. When implied volatilities are calculated, the life of an option should be measured in trading days. In the chapter, daily data are used to provide a historical volatility estimate where we ignore days when exchange is closed. Usually, implied volatilities are scaled as annualized standard deviations.

4.2.2 Empirical data

We consider the closing prices of European call options written on the S&P500 index with strike prices ranging from 775 through 1565 index points, traded in the Chicago Board of Options Exchange (CBOE). At-the-Money (ATM) options are employed with a constant time to maturity of 30 days. The risk free rate is the secondary market 3-month U.S. treasury bill rate. The option data, C_{market}, the S&P500 index observations, S, and the secondary market 3-month U.S. treasury bill rates, r_f, were collected from Datastream[4] for the period from 15 January 2002 till 14 March 2008. Totally, there are 1609 daily observations for each time series.

The daily returns, R_t, of call options on the S&P 500 index are changes in the logarithms of daily option prices P_t as follows

$$R_t = \ln(\frac{P_t}{P_{t-1}}) * 100 \tag{4.3}$$

where the time counter t refers to trading days. The excess return to the risky asset is

$$\varrho_t \equiv R_t - r_f \tag{4.4}$$

We use ϱ_t to forecast one-step ahead conditional variance of the properly aggregated excess return series. Daily squared excess return ϱ_t^2 is used to proxy daily realized volatility[5] $\sigma_{R,d}^2$

$$\sigma_{R,d}^2 = \varrho_t^2 \tag{4.5}$$

where ϱ_t^2 is an unbiased estimator of $\sigma_{R,d}^2$. See Poon and Granger (2003) and Lopez (2001) for the detailed discussions.

We divide the call option data into two subsamples: the most recent 522 daily observations (from 16 Mar 2006 to 14 Mar 2008) are treated as "out-of-sample" data for forecast while the rest 1087 daily observations (from 15 Jan 2002 to 15 Mar 2006) are the "in-sample" data for estimation. The annualized realized volatility is

$$\sigma_{R,a} = \sqrt{N}\sigma_{R,d} \tag{4.6}$$

[4]The empirical data were collected from Thomson Datastream (4.0 version). The secondary market 3-month U.S. treasury bill rate is middle rate. Treasury bills are short-term securities issued by the U.S. Treasury. Treasury Bills are traded in primary and secondary markets. Secondary trading in Treasuries occurs in the over-the-counter (OTC) market. Ratea are annualized using a 360-day year or bank interest on a discount basis. The rest corresponding data of X and q to the Black-Scholes pricing formulae are available as well from Datastream for the same time period. Given the known C_{market}, S, X, T, r_f, and q, a backward induction technique in Eq. (4.1) can be used to derive $\sigma_{implied}$.

[5]Before high frequency data becomes widely available, many researchers have resorted to using daily squared return, calculated from market closing prices, to proxy daily volatility. Given that volatility is a latent variable, the actual volatility is often estimated from a sample. For a long time, σ_t is proxied by daily squared return if t is a day. For the high frequency data, daily σ_t is derived from the cumulation of intra-day returns. See Poon and Granger (2003) for the details.

where N denotes the number of trading periods in one year. Typically, there are 252 trading days in one year.

The statistical characteristics of the distributions of the daily call option and S&P 500 index returns are summarized in Table 4.1. The statistics show that the distributions of both time series are not normal. In details, each distribution has high kurtosis, fat tails and a peaked centre compared with the normal distribution. The hypothesis of the normal distribution is rejected at the 1% level for both series.

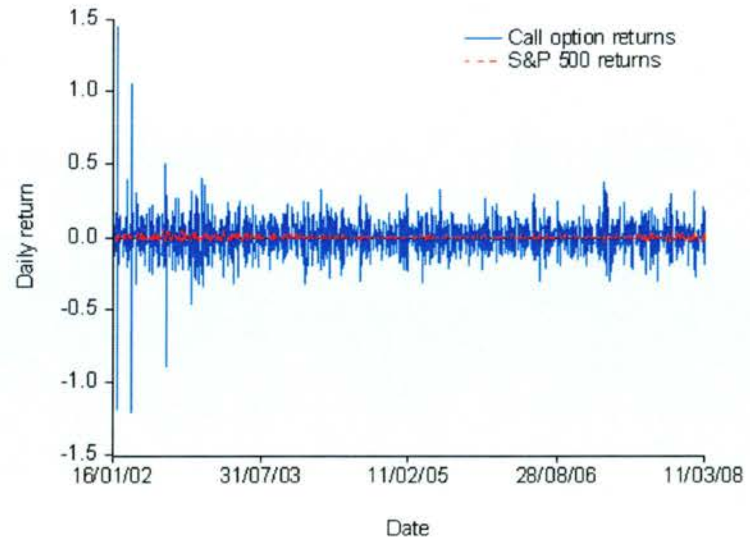
Figure 4.1 plots "volatility clustering" in the daily returns of both call options and S&P 500 index for January 2002 - March 2008. The current level of volatility tends to be positively correlated with its level during the immediately preceding periods. The important point to note from Figure 4.1 is that volatility occurs in bursts. There appears to have been a prolonged period of relative tranquility in the market during the mid-2000s, evidenced by only relatively small positive and negative returns. On the other hand, roughly during early 2002 to mid-2003 ("dot-com bubble" of the early 2000s recession) and mid-2007 to earlier 2008 ("subprime mortgage crisis"), there was far more volatility, when many large positive and large negative returns were observed during a short space of time. It demonstrates that volatility is autocorrelated.

As shown, the statistics in Table 4.1 and the patterns in Figure 4.1 illustrate that changes in call option prices derived from the underlying S&P 500 index are more volatile than those of the S&P 500 index levels itself. Call option tends to be more risky than S&P 500 index for the period of January 2002 to March 2008. The portfolio of call option on the S&P 500 returns we use can be better diversified.

Return Data	Mean	Median	Max.	Min.	Std. Dev.	Skewness	Kurtosis	Jarque-Bera	Probability
Call option	0.000	0.000	1.438	-1.205	0.125	0.085	28.638	44040.40	0.000
S&P 500 index	0.000	0.000	0.056	-0.042	0.010	0.090	5.834	540.0974	0.000

Notes: European call option is written on the S&P 500 index. Std. Dev. (standard deviation) is a measure of dispersion or dispersion or spread in the series. Skewness is a measure of asymmetry of the distribution of the series around its mean. Kurtosis measures the peakedness or flatness of the distribution of the series. Jarque-Bera test statistic measures the difference of the skewness and kurtosis of the series with those from the normal distribution. The reported probability is the probability that a Jarque-Bera statistic exceeds (in absolute value) the observed value under the null hypothesis of a normal distribution - a small probability value leads to the rejection of the null hypothesis. The sample size is 1608 daily observations of log returns as defined in Equation (4.3) for the period from 16 January 2002 to 14 March 2008

Table 4.1: Summary statistics for time series of daily call option and SP500 index returns



Notes: European call option is written on the S&P 500 index.

Figure 4.1: Daily call option and SP500 index returns for January 2002 - March 2008

4.2.3 Volatility models and estimation techniques

A variety of techniques have been developed to forecast volatility in financial markets. Most fall into econometric and statistical techniques[6]. Taking account of the stylized fact that for many financial variables, including equity options, squared changes display an important feature of "volatility clustering", in the chapter, we focus on the most popular autoregressive conditionally heteroscedastic (ARCH) models that provide a vast variety of volatility forecasts.

In contrast to historical volatility models[7], ARCH class models do not make use of sample standard deviation, but formulate conditional variance, h_t^2, of returns via maximum likelihood procedure which works by choosing coefficient estimates that maximize

[6]The chapter excludes discussions of volatility models basing on some techniques associated with neural networks, genetic programming, time change and duration, signal processing, and spectrum analysis etc.

[7]The group of historical volatility models includes random walk, historical average, autoregressive moving average, and various forms of exponential smoothing that depend on the values of the weight parameter. The key feature for this type of the models is that predictions are based on past standard deviations. See Chapter 2 for the details.

the likelihood of the actual sample data set. As discussed in Poon and Granger (2003), given the construction of ARCH class models where h_t^2 is known at time $t-1$, the one-step ahead forecast is readily available, and the forecasts that are more than one step ahead can be formulated based on an iterative procedure. For more surveys, see Bera and Higgins (1993), Bollerslev, Chou, and Kroner (1992), Bollerslev, Engle, and Nelson (1994), and Diebold and Lopez (1995).

Assuming that the return series of a financial asset, r_t, has the following stochastic process

$$r_t = \mu_t + \varepsilon_t$$

$$\varepsilon_t = \sqrt{h_t^2} z_t \qquad z_t \sim (0,1)$$

where $E(r_t \mid I_{t-1}) = \mu_t$, $E(\varepsilon_t^2 \mid I_{t-1}) = h_t^2$ given the information set I_{t-1} at time $t-1$, and z_t is conditionally normally distributed with zero conditional mean and unit conditional variance so that ε_t will also be conditionally normally distributed with zero conditional mean and conditional variance h_t^2.

In the chapter, we consider a symmetric generalized ARCH (GARCH) model of Bollerslev (1986) and Taylor (1986), an asymmetric exponential GARCH (EGARCH) model of Nelson (1991), and an asymmetric component GARCH (CGARCH) model. For GARCH(1,1) model, h_t^2 follows

$$h_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}^2 \tag{4.7}$$

where $\omega > 0$. See Nelson and Cao (1992) for constrains on α and β in details. For finite variance, $\alpha + \beta < 1$. It is noted that EWMA can be viewed as a non-stationary version of GARCH(1,1) (see, for example, Poon and Granger, 2003 and González-Rivera, Lee and Mishra, 2004).

For EGARCH(1,1) model, h_t^2 follows

$$\ln(h_t^2) = \omega + \beta \ln(h_{t-1}^2) + \alpha \left[\left| \frac{\varepsilon_{t-1}}{h_{t-1}} \right| - \sqrt{\frac{2}{\pi}} \right] + \boldsymbol{\eta} \frac{\varepsilon_{t-1}}{|h_{t-1}|} \tag{4.8}$$

For CGARCH(1,1) model, h_t^2 follows

$$h_t^2 = \omega + \rho\left(q_{t-1} - \omega\right) + \varphi\left(\varepsilon_{t-1}^2 - h_{t-1}^2\right) + \alpha\left(\varepsilon_{t-1}^2 - q_{t-1}\right) + \gamma\left(\varepsilon_{t-1}^2 - q_{t-1}\right) D_{t-1} + \beta\left(h_{t-1}^2 - q_{t-1}\right) \tag{4.9}$$

In the chapter, we use GARCH(1,1), EGARCH(1,1) and CGARCH(1,1) models for one-step ahead daily volatility forecast since day t, respectively. $\widehat{\sigma}_{F,d}$ is the forecasted daily conditional volatility of the actual daily volatility $\sigma_{R,d}$, $\widehat{\sigma}_{F,d} = \sqrt{\widehat{h}_t^2}$ where $\widehat{h}_t^2$ is the forecasted conditional variance. All parameters are reestimated under a rolling window where the start date and end date successively increase by one observation every time. $\widehat{\sigma}_{F,d}$ in turn will be used to estimate the price of the call option.

4.3 Expected utility maximization

It is a one period problem where an investor is assumed to know conditional excess returns and cares about getting the volatility forecast right because it will affect his optimal desired holdings of risky assets. Effectively it is saying that getting good forecasts of excess returns is not the main problem for the investor. Rather, getting good forecasts of volatility to enable the investor to hold the "right" amounts of the risky asset is assumed to be the main worry for the investor. This is a credible story for foreign exchange rates and call options on the S&P index that have extreme time varying volatility in their excess returns but where the expected excess returns themselves are well known in the financial practitioner world. Importantly, González-Rivera, Lee and Mishra (2004) use the S&P 500 index. In a one factor model like CCAPM/ICAPM returns on this index are the sole source of undiversifiable risk. Hence it make sense that the risk in the portfolio is all that matters for utility. Put another way, the covariance of the portfolio with the market (the beta) is equal to the alpha times the market portfolio's variance. Hence using a single asset for this exercise is inappropriate because although a single asset will have non zero variance, much of the risk associated with it may be completely diversified and hence will not be priced and there will not therefore be a close link between asset return variance and utility. In the chapter, we consider a portfolio that consists of a risky asset (e.g. European call option written on the S&P 500 index) and a riskless asset (e.g. the 3-month treasury bill).

Given different correlated assets, how does an investor decide asset demand and create a portfolio maximizing the expected utility? An investor might do well, relying only on the means and variances/covariances of the asset returns, which simplifies the portfolio selection tremendously. The validity of the mean-variance (MV) approximation to exact utility maximization has been verified in the case of choosing securities. Returns on the mean-variance frontier can be generated as portfolios of any two frontier returns, where all assets returns lie inside a mean-variance frontier and assets on the frontier are perfectly correlated with each other and with the discount factor.

The mean-variance analysis developed by Markowitz critically relies on two assumptions: either the investors have quadratic utility or the asset returns are jointly normally distributed. There is no need for both assumptions, just one or the other is required: (1) If an investor has quadratic preferences, he cares only about the mean and variance of returns; and the skewness and kurtosis of returns have no effect on expected utility, i.e., he will not care, for example, about extreme losses. Quadratic utility has been shown to be inconsistent with observed human choice behaviour with respect to risk. (2) Mean-variance optimization can be justified if the asset returns are jointly normally distributed since the mean and variance will completely describe the distribution. The normal distribution is symmetric, thus its skewness or third moment is zero. The kurtosis or fourth moment of a normal distribution has a value equal to three. A number of researchers have asserted that the right choice of mean-variance efficient portfolio will give precisely optimum expected utility if and only if all distributions are normal or investors have quadratic utility function.

In the chapter, we consider portfolio optimization of a two-security case in an extended Markowitz framework. We assume that the investor maximizes his expected utility given that his wealth is allocated between a risky asset with random return and a riskless asset with sure return where his coefficient of relative risk aversion is constant. In particular, we reflect on an individual investor who maximizes his single-period expected utility. The corresponding preference function is expressed solely in terms of expected return and variance, which can be justified by either a quadratic utility function or normally distributed asset returns (see Haugen, 2003, page 201-204). We also assume that all relevant moments of the return distribution are known except for the conditional variance. The investor must try and find an estimate of forecasted conditional variance in order to work out his best guess of the optimal proportion of wealth placed in risky assets. If he underestimates the conditional variance he will probably hold too much of the asset (too much relative to the the proportion of wealth placed in risky assets that would maximize utility) and vice versa for overestimates. Note that wealth does not enter the utility function. This is because in a one period static problem, end period wealth is just (one plus) the return on the portfolio times initial wealth which is a known constant. Therefore maximizing one step ahead expected utility amounts to maximizing one step ahead expected utility of returns. In the studies of asset demand, we will derive maximized utility from two special cases – either a quadratic utility function (see Tobin, 1958) or a negative exponential utility function (see Freund, 1956 and Parkin, 1970) with the assumption that the probability distributions for returns are normally distributed, (see Bhattacharyya, 1979).

Before proceeding to the derivations, we list the justifications for choosing these special cases, as well as the assumptions and implications of relevant decision making. The quadratic utility function has traditionally occupied a place of importance both in financial analysis and in academic expositions of financial decision making under uncertainty. Its appeal, as introduced in Gregory (1978), is due in large measure, to "its simplicity and mathematical tractability, as well as to its ability to serve as an approximation (second order) to more complicated utility functions". The primary theoretical objection to quadratic utility is that an increase in assets results in both increased absolute and relative risk aversions, for example, the certainty equivalent of dollars and percentage of wealth investors are willing to commit to risky investments respectively, in the sense that investors' willingness to take on risk decreases as wealth level increases. Pratt (1964) labeled this property of risk preference to situations in which less wealth will be preferred to more wealth.

The second case considered assumes that the decision maker's risk preference is representable by an exponential utility function. The exponential utility function is frequently used in applied decision-analytical work (e.g. Keeney and Raiffa (1976), p211) and treated as an applied tool (e.g. Hammond (1974), p1056). The exponential utility function is constantly absolutely risk averse (CARA)[8]. Therefore, increases in assets do not affect the certainty equivalent for a fixed risk, a property which may or may not be reasonable depending on the decision context. The normal distribution of returns is combined with the exponential utility function where the normal distribution

[8]The exponential utility function is constantly absolutely risk averse since the Pratt-Arrow risk index is constant. The quadratic utility function is increasingly absolute risk aversion since the Pratt-Arrow risk index is increasing.

has frequently been used to represent cash flow distributions because of its tractability and statistical properties.

4.3.1 Quadratic utility function

Following González-Rivera, Lee and Mishra (2004), the quadratic time-independant additive utility function is

$$U_{t+1} = w_{t+1} - 0.5\gamma w_{t+1}^2 \tag{4.10}$$

with wealth constraint

$$w_{t+1} = \alpha_t r_{t+1} + (1 - \alpha_t) r_{f,t+1} \tag{4.11}$$

where U_{t+1} denotes utility at time $t + 1$, w_{t+1} denotes the return to the portfolio at time $t + 1$, γ denotes the risk aversion parameter, α_t denotes the proportion of the risky asset held at time $t + 1$, r_{t+1} denotes the gross return on the risky asset at time $t + 1$, and $r_{f,t+1}$ denotes the risk free rate of interest for one period which is assumed to be constant ($r_f = r_{f,t+1}$). Initial wealth W_0 is assumed to equal to one. Hence the portfolio's conditional variance is $\alpha^2\sigma^2_{r,t+1}$[9].

The expected quadratic utility[10] of Equation (4.10) is

$$E_t(U_{t+1}) = E_t\left[w_{t+1} - 0.5\gamma w_{t+1}^2\right] \tag{4.12}$$

where E_t denotes the mathematical conditional expectations operator at time $t + 1$. Therefore, for a one-period horizon, the maximized expected quadratic utility in Equa-

[9] $Var[\alpha r_{t+1} + (1-\alpha)r_{f,t+1}] = \alpha^2 Var(r_{t+1}) + (1-\alpha)^2 Var(r_{f,t+1}) + 2\alpha(1-\alpha)Cov(r_{t+1}, r_{f,t+1})$ where r_{t+1} and $r_{f,t+1}$ are not correlated so that $Var[w_{t+1}] = Var[\alpha r_{t+1} + (1-\alpha)r_{f,t+1}]$ $=\alpha^2 Var(r_{t+1})$. In other words, $\sigma_{w,t+1} = \alpha^2\sigma_{r,t+1}$.

[10] The expected value of the quadratic function of portfolio return in Eq. (4.10) can be written as $E_t(U_{t+1}) = E_t(w_{t+1}) - 0.5\gamma\left[\sigma^2_{w,t+1} + E_t(w_{t+1})^2\right]$ where $\sigma^2_{w,t+1}$ denotes the conditional variance of the portfolio returns. Thus the expected quadratic utility is defined in terms of means and variances. We consider a mean-variance maximizing investor who maximizes the quadratic utility function.

tion (4.12) with respect to α_t subject to Equation (4.11) is

$$\max_{\alpha_t} E_t(U_{t+1}) \equiv E(w_{t+1} - 0.5\gamma w_{t+1}^2) \tag{4.13}$$

Equation (4.13) leads to get the investor's optimal α_t. This will be a function of $r_{f,t+1}$, γ, μ_{t+1} and $\widehat{\sigma}_{t+1}^2$. From the first order conditions, we obtain the optimal portfolio weight of the risky asset as follows

$$\alpha_t = \frac{\mu_{t+1}(1 - \gamma r_{f,t+1})}{\gamma(\mu_{t+1}^2 + \widehat{\sigma}_{t+1}^2)} \tag{4.14}$$

where $\varrho_{t+1} \equiv r_{t+1} - r_{f,t+1}$ is the excess return to the risky asset, μ is the expected excess return on the risky asset, $\mu_{t+1} \equiv E(\varrho_{t+1})$, σ_{t+1}^2 is the true conditional variance of the conditional expected asset return ϱ_{t+1}, $\widehat{\sigma}_{t+1}^2$ is the estimated conditional variance of σ_{t+1}^2 from the forecasting models (e.g. GARCH). All expected returns and variances are conditional on information at time t. σ_{t+1}^2 is itself a forecast - it is in fact the rational expectations one step ahead forecast of one step ahead returns using the true parameter values of the model. Investors in the model are assumed not to know the true parameters of the model – if the investor was endowed with rational expectations in the strict sense of actually knowing all the model's parameters, particularly σ_{t+1}^2, he could choose α_t to maximize this - but he does not have rational expectations because although he knows $r_{f,t+1}$, γ, and μ_{t+1} he does not know the true forecasted volatility σ_{t+1}^2 but must use an estimate of this from a forecast model. We have called this estimate $\widehat{\sigma}_{t+1}^2$. The investor must try and find an estimate of $\widehat{\sigma}_{t+1}^2$ in order to work out their best guess of optimal α_t. If they underestimate the conditional variance they will probably hold too much of the asset (too much relative to the α_t that would maximize utility) and vice versa for overestimates. The formula in Equation (4.14) is the best the investor can do to choose optimal α_t.

We use the recursive estimation approach in Section 4.2.3 to obtain forecasts of the conditional variance, $\widehat{\sigma}^2_{t+1}$. These estimates can be used in Equation (4.14) to compute the conditional expectations and conditional variance of wealth at time $t+1$. Specifically, for each period, we use a model of interest to choose the fraction of wealth that maximizes expected utility, taking the model's point estimate for the conditional variance as the correct expectation. Given the assumption that the mean return on the asset is known, in a given period, the optimal proportion will vary across competing models only insofar as the estimates of the conditional variance vary. Hence, we rewrite $E_t(U_{t+1})$ now replacing α_t of $\widehat{\alpha}_t$ from Equation (4.14) and the investor's mathematically expected utility may then be written

$$E_t(\widehat{U}_{t+1}) = c_{t+1} + \zeta_{t+1}\varkappa(\sigma^2_{t+1}, \widehat{\sigma}^2_{t+1}) \tag{4.15}$$

where

$$c_{t+1} \equiv r_{f,t+1} - 0.5\gamma r^2_{f,t+1} \tag{4.16}$$

$$\zeta_{t+1} \equiv \mu^2_{t+1}\frac{(1-\gamma r_{f,t+1})^2}{\gamma} \tag{4.17}$$

and

$$\varkappa(\sigma^2_{t+1}, \widehat{\sigma}^2_{t+1}) \equiv \frac{1}{\mu^2_{t+1} + \widehat{\sigma}^2_{t+1}} - 0.5\frac{\mu^2_{t+1} + \sigma^2_{t+1}}{(\mu^2_{t+1} + \widehat{\sigma}^2_{t+1})^2} \tag{4.18}$$

The expression is the investor's maximized utility and it depends on $r_{f,t+1}$, γ, μ_{t+1}, σ^2_{t+1} and $\widehat{\sigma}^2_{t+1}$. We call this expression estimated expected utility $E_t(\widehat{U}_{t+1})$. Because the investor does not know σ^2_{t+1}, the result given in Equation (4.15) cannot directly used to evaluate forecasts of a risky asset's volatility. However, West, Edison and Cho (1993) have proposed to get an estimate of σ^2_{t+1} that is right on average by substituting the ex-post realized squared excess return to the risky asset ϱ^2_{t+1} for its conditional expectation σ^2_{t+1}. Hence, Equation (4.18) may be written

$$\varkappa(\varrho^2_{t+1}, \widehat{\sigma}^2_{t+1}) \equiv (\mu^2_{t+1} + \widehat{\sigma}^2_{t+1})^{-1} - 0.5(\mu^2_{t+1} + \varrho^2_{t+1})(\mu^2_{t+1} + \widehat{\sigma}^2_{t+1})^{-2} \tag{4.19}$$

Investors's expected utility can be rewritten as

$$E_t(\widehat{U}_{t+1}) = c_{t+1} + \widehat{\zeta}_{t+1}(\gamma)\,\widehat{\varkappa}(\varrho^2_{t+1}, \widehat{\sigma}^2_{t+1}) \tag{4.20}$$

where $\widehat{\zeta}\,(.)$ and $\widehat{\varkappa}(.,.)$ are obtained from Equations (4.17) and (4.19) by replacing μ_{t+1} with the predicted excess return $\widehat{\mu}_{t+1}$ (see González-Rivera, Lee and Mishra 2004). It is emphasized that wealth does not enter the U function. This is because in a one period static problem, end period wealth is just (one plus) the return on the portfolio times initial wealth W_0 which is a known constant. Therefore maximizing one step ahead expected utility amounts to maximizing one step ahead expected utility of returns. In other words, if one is to keep a constant amount of wealth W_0, one's consumption must be equal to rW_0 where r is the current portfolio return. Consume more than this and the investors must run down wealth - consume less and the investors build wealth. Explicit formulas are given in Appendix C.1.

We consider to choose the values for the risk aversion parameter γ. In the empirical section, we set γ at 0.1, 0.3, 0.5, 0.7, and 0.9. The reason is that, thinking of the quadratic utility function as a second order approximation to the power utility function $U = \frac{w^{1-\rho}}{1-\rho}$, $\gamma = \frac{\rho}{1+\rho}$[11]. Being familiar with power utility from Moore and Roche (2006) studied in the previous chapter, where two values of ρ (and hence two values of γ) that we think empirically reasonable and Moore and Roche would be happy to use in their model. Specifically, Moore and Roche (2006) used $\rho = 0.5$ and $\rho = 0.7$ in the baseline and sensitivity analysis so that the corresponding γ values calculated are $\gamma \approx 0.33$ and $\gamma \approx 0.41$ respectively, which is the benchmark for us to choose the risk aversion coefficient. However, in order to investigate changes when γ varies, we extend

[11]For the power utility function, the coefficient multiplying the quadratic term is (normalizing the coefficient on the linear term to unity and ignoring constants) $-\frac{1}{2}\frac{\rho}{1+\rho}$. Equate this with $-\frac{\gamma}{2}$ in the quadratic utility function. In other words, let $\rho \equiv \gamma W/(1-\gamma W)$ where $W = 1$ for simplicity be the coefficient of relative risk aversion. Hence, $\gamma = \frac{\rho}{1+\rho}$.

this benchmark to a wider range of the values of γ on the interval of $[0.1, 0.9]$. We find that the results are robust when we have experimented with different values of the risk aversion coefficient.

4.3.2 Exponential utility function and normally distributed returns

Following Freund (1956), Chopra and Ziemba (1993) and others, we assume a negative exponential utility function

$$u(w_{t+1}) = -e^{-Aw_{t+1}}, \quad A > 0 \tag{4.21}$$

where w_{t+1} is a measure of the portfolio return[12] at time $t+1$ and A the coefficient of absolute risk aversion. The larger the value of A, the more conservative the investor. If w_{t+1} is normally distributed then $-Aw_{t+1}$ is also normal and hence $-E_t(e^{-Aw_{t+1}})$ is the negative of the moment-generating function (MGF) of the normal variable $-Aw_{t+1}$, we have

$$E_t[u(w_{t+1})] = -E_t(e^{-Aw_{t+1}}) = -e^{-A\mu_{w_{t+1}}+(A^2/2)\sigma^2_{w_{t+1}}} \tag{4.22}$$

where $\mu_w \equiv E(w)$ and $\sigma_w^2 = E\left[(w-\mu_w)^2\right]$ are the expected return and variance of the portfolio. The combination of exponential utility and normal distribution has a particularly convenient analytical form. Since it gives rise to linear demand curves, it is very widely used in models that complicate the trading structure, by introducing incomplete markets or asymmetric information. For example, Grossman and Stiglitz (1980) is a very famous example. Normally distributed returns make this consistent with the mean-variance framework. The negative exponential utility function is especially convenient in a world of normally-distributed outcomes. Recall that expected

[12]In standard utility theory the argument is the absolute value of wealth at a future date. Some assume that such a function can be applied repeatedly for one-period decisions on sequential dates. However, for purposes of portfolio theory it is desirable to state utility in terms of return (the relative change in wealth over the future period). In the chapter, we present a model with return w only in the last period for one-period decisions.

utility is the integral of the utility function using the probability distribution as weights. If the former is negative exponential and the latter is normal, it will be the case that expected utility will be a simple function of the mean and variance of the distribution. For example, if return w is normally distributed then $-Aw$ is also normal and hence $-E(e^{-Aw})$ is the negative of the moment generating function (mgf) of the normal variable $-Aw$ and hence $-E(e^{-Aw}) = -e^{-AE(w)+(A^2/2)var(w)}$. See Sargent (1987), page 154, for the details.

Suppose this investor has terminal return (at time $t+1$) of the portfolio consisting of a risk-free asset paying $r_{f,t+1}$ and a risky asset paying r_{t+1}. Let g_t denote the weight of the risky asset in the portfolio. Plugging the budget constraint $w_{t+1} = g_t r_{t+1} + (1-g_t) r_{f,t+1}$ into the exponential utility function, we obtain

$$E_t u\left(w_{t+1}\right) = -e^{-A\left[g_t E(r_{t+1}) + (1-g_t) r_{f,t+1}\right] + (A^2/2) g_t^2 \left[E(r_{t+1}^2) - (E(r_{t+1}))^2\right]} \tag{4.23}$$

Maximizing Equation (4.23) with respect to g_t, we obtain the first-order condition describing the optimal fraction of the risky asset held at time $t+1$,

$$g_t = \frac{\mu_{t+1}}{A\widehat{\sigma}_{t+1}^2} \tag{4.24}$$

where $\widehat{\sigma}_{r,t+1}^2$ is the estimated variance of risky returns.

$$E_t(\widehat{u}_{t+1}) = -e^{-\mu_{t+1}^2/\widehat{\sigma}_{t+1}^2 - Ar_{f,t+1} + \frac{1}{2}\sigma_{t+1}^2 \frac{\mu_{t+1}^2}{(\widehat{\sigma}_{t+1}^2)^2}} \tag{4.25}$$

See Appendix C.2 for the derivation details. We have chosen and experimented with different values of A on the interval of $[0.1, 0.9]$, which are the same as those to γ in the quadratic utility function. Our results remain unchanged.

Sensibly, the investor invests more in risky assets if his expected return is higher, less if his risk aversion coefficient is higher, and less if the assets are riskier. Notice that total wealth does not appear in this expression. With this setup, the amount invested in risky assets is independent of the level of wealth. This is why it is said that this investor has absolute rather than relative (to wealth) risk aversion. Note also that these "demands" for the risky assets are linear in expected returns.

4.3.3 Asymmetry of utility function

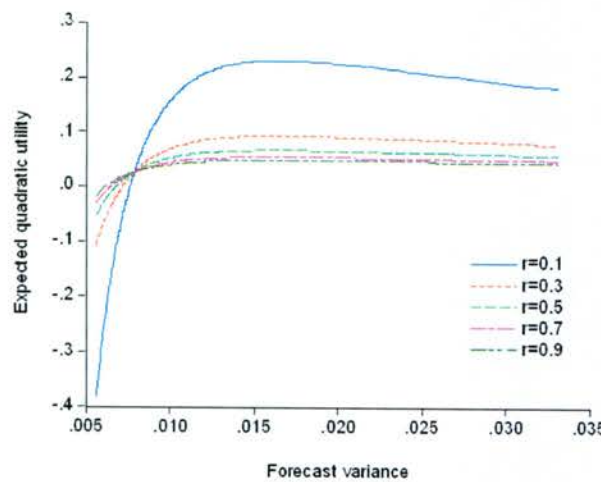
The quadratic and exponential utility functions are asymmetric. Both of them depend on the utility risk aversion parameter, one-step-ahead ex post variance, and one-step-ahead forecast variance. Miscalculations of the conditional variance are paid in units of utility.

To illustrate the asymmetry, we consider a graph, which plots expected utility as a function of $\widehat{\sigma}^2_{t+1}$, the estimate of σ^2_{t+1}, when σ^2_{t+1} is a scalar and utility is either quadratic (i.g. $E_t(\widehat{U}_{t+1})$) or exponential (i.g. $E_t(\widehat{u}_{t+1})$) with parameters matching those in our empirical work. By assumption, highest expected utility occurs when $\widehat{\sigma}^2_{t+1} = \sigma^2_{t+1}$, which is the maximum and is the peak. In details, using empirical data averages for $r_{f,t+1}$, μ_{t+1}, and σ^2_{t+1} we graph utility functions $E_t(\widehat{U}_{t+1})$ and $E_t(\widehat{u}_{t+1})$ on the vertical axis against values of the forecasts of one step ahead conditional variance of the conditional expected excess asset return $\widehat{\sigma}^2_{t+1}$ respectively for some true value of one step ahead conditional variance (taken from the data itself). The graphs in Figure 4.2 obviously have a peak and show a maximum exactly at the point on the x-axis when forecast equals true value (forecast error is zero) where $\widehat{\sigma}^2_{t+1} = \sigma^2_{t+1} = (0.128)^2 = 0.016$. The peaks are in exactly the same place for both the objective functions of utility-based criteria. As displayed in Panel A for quadratic utility function and in Panel B for exponential utility function of Figure 4.2, it is found that expected utility declines the farther away $\widehat{\sigma}^2_{t+1}$ is from σ^2_{t+1}: for values of $\widehat{\sigma}^2_{t+1}$ below (i.e. to the left of) σ^2_{t+1}, it is seen that utility is a lot lower (steep slope) but to the right of σ^2_{t+1} is only a little bit lower (shallow slope). They show how quickly or slowly utility changes when the forecasted conditional variance is on either side of the true value – indicating for example the importance of positive versus negative forecast errors. In contrast to the usual mean squared or mean absolute error criterion, both the utility functions are asymmetric around σ^2_{t+1}, penalizing estimates that are too small more sharply than those that are too large. This is telling us that mistakes in estimating conditional volatility are more serious when σ^2_{t+1} is underestimated than when it is overestimated. The results remain unchanged when we have chosen and experimented with different values of the risk aversion coefficient at 0.1, 0.3, 0.5, 0.7, and 0.9. As we shall see, this asymmetry plays a role in the empirical results.

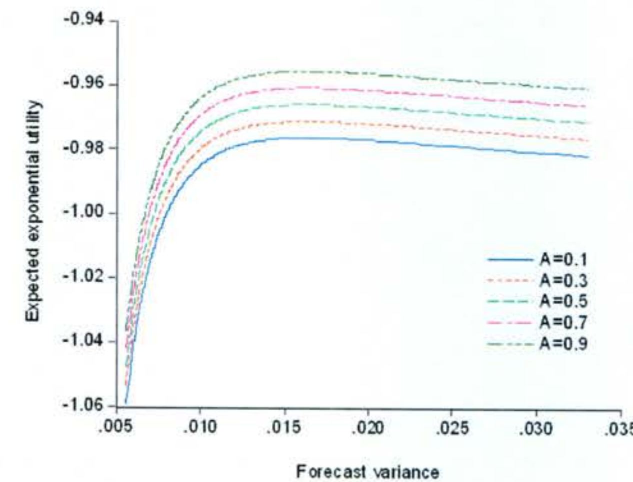
4.4 Optimal forecast error criteria

We aim to find an approximation to the relationship in which expected utility is a function of forecast error. The utility-based economic loss functions in Equations (4.20) and (4.25) are in terms of call option price forecast errors which – via Equation (4.1) – depend on volatility forecast errors and hence these errors enter in a nonlinear and analytically intractable way into the loss functions. In general, forecast errors will affect the value of expected utility of wealth. The functional form could be numerically established with the mapping from errors to wealth under the trading rule. In the sense that making forecast errors should be bad for utility, the expected utility should be a declining function of forecast error etc. Using these utility-based economic loss functions may well lead to high profits in trading but no econometrician or statistician would recognize it or be able to analyze its statistical properties. We will try to find some sort of analytical loss function, e.g. linear combination of statistical MAE or MSE etc., that would lead us to select a volatility model which is very close to that

Panel A:
Graph of $E_t(\widehat{U}_{t+1})$ as a function of $\widehat{\sigma}^2_{t+1}$, $\sigma^2_{t+1} = 0.016$



Panel B:
Graph of $E_t(\widehat{u}_{t+1})$ as a function of $\widehat{\sigma}^2_{t+1}$, $\sigma^2_{t+1} = 0.016$



Notes: In Panel A, r denotes the risk aversion parameter in quadratic utility function. In Panel B, A denotes the risk aversion parameter in exponential utility function.

Figure 4.2: Asymmetry of expected utility at different values of the risk aversion coefficient

we get from using economic loss function with similar loss.

The above analysis in Section 4.3.3 will show that for these types of utility functions negative ARCH forecast errors are less costly than positive ones. Note that because the problem is static there is no connection between utility at time $t+1$ and errors made before or after time $t+1$. Only time $t+1$ errors affect utility at time $t+1$. Hence we wish to find the the function f such that

$$\widetilde{U}_{t+1} \simeq f(e_{t+1}) \tag{4.26}$$

where e_{t+1} is the one step ahead ARCH model forecast error, $e_{t+1} = \sigma_{t+1} - \widehat{\sigma}_{t+1}$. The graphs in Figure 4.2 in the previous section suggest that f is asymmetric so for example if f was linear positive $e's$ should enter with a separate coefficient to negative ones. However f of course will not be linear so the problem is how to "estimate" f.

In our empirical work, for out-of-sample forecasts of conditional volatility, we have about 522 trading days (from 16 Mar 2006 to 14 Mar 2008) of one-step-ahead forecast errors and corresponding utility values (we did quadratic and exponential utility cases) for each of three ARCH models. This makes $1566 (= 3 * 522)$ pairs of $\left\{\widetilde{U}_i, e_i\right\}$ for use in regression analysis to try and discover the form of f. We suggest the following regression using these pairs as data. We create a dummy variable d_i for each $\left\{\widetilde{U}_i, e_i\right\}$ pair that is zero if e_i is positive and -1 if e_i is negative. Then we regress

$$\widetilde{U}_i = \beta_0 + \beta_1 |e_i| + \beta_2 d_i e_i + \beta_3 e_i^2 \tag{4.27}$$

It is emphasized that β is determined by the form of the utility function and the risk aversion coefficient but not by the specific model (e.g. GARCh or EGARCH or CGARCH etc.) being used. An investor would expect that whatever model he uses (e.g. GARCh or EGARCH or CGARCH etc.) should give the same β for a given risk aversion coefficient. Roughly, we see that that is the case in our estimates, which shows that it is working well. So for each risk aversion coefficient (and for each of exponential and quadratic utility functions separately) we should "pool" all observations together to get maximum efficiency for the regression. For example, if we have 10 risk aversion coefficients we do 10 regressions for quadratic utility case and then 10 for exponential utility case.

Averaging left and right hand sides will give us a statistical criterion C that if maximized will give us optimum average utility. It will be

$$C = \beta_1 MAE + \beta_2 MANE + \beta_3 MSE \tag{4.28}$$

where MAE is mean absolute error, $MANE$ is the mean absolute of the negative error, and MSE is mean squared error. In the chapter, the accuracy of forecasts is evaluated by loss function. For statistical convenience, we consider symmetric MAE and MSE of statistical loss functions that are some simple function of the forecast error. $MAE = N^{-1} \sum_{N=1}^{1566} |e_i|$ and $MSE = N^{-1} \sum_{N=1}^{1566} e_i^2$, where N is the out-of-sample size. The regression goes through the mean because it has an intercept. When we take averages of all sides we get average utility as well as different statistical loss functions. This averaged form is the approximate optimal forecast error criterion for this particular investor facing this particular investment decision. Note that Equation (4.28) is simply a rewriting of the regression in Equation (4.27) in mean form where all regressions with intercepts go through the mean.

As shown in Figure 4.2 in Section 4.3.3, the shape of the plots we draw seems to be steep and nearly linear to the left (the "steep" left in short) and shallow and slightly convex to the right (the "shallow" right in short). So it should have (as an approximation) low but positive β_3, negative β_2 and negative β_1. Specifically, for either of the quadratic and exponential utility cases, if the objective function has a peak when forecast variance equals to expected/real variance, $\widehat{\sigma}_{t+1}^2 = \sigma_{t+1}^2$, then the error is zero.

Utility starts to decrease when (positive and negative) error exists. In this case, the linear effect of positive errors is $(\beta_0 + \beta_1 + \beta_3)$ and the linear effect of negative errors is $(\beta_0 + \beta_1 + \beta_2 + \beta_3)$. The quadratic effect[13] (curvature) is represented by β_3. We can plot a tangent line[14] to the "steep" curve on the left of the peak(s) that passes through the "hump" point. We find, but do not report, that the tangent line is a upward straight-line with a low slope. β_3 is not greater than the slope of the tangent line at the "hump" point. So β_3 is positive but low. Furthermore, the area between the tangent line and the "steep" curve on the left of the peak(s) is smaller than that between the tangent line and the "shallow" curve on the right of the peak(s) so that $\beta_1 < 0$ and $\beta_2 < 0$[15]. This is what our graphs in Figure 4.2 tell us: positive errors (or underestimates) lead to lower utility than negative errors (or overestimates) do. What is more, the squared error term to β_3 is like a squared effect on the graph - it becomes dominant for large errors so a positive effect of squared error will pick up the upward ski-slope shape appearing on the far right of the peak(s). Quite probably there is too little room between the vertical axis and the peak for the squared error term to have any meaningful impact in term, hence the ski-slope upturn.

Of course, the β coefficients will depend on the utility function and extent of risk aversion in it. We repeat the exercises for different values of risk aversion and of course for the two utility functions - quadratic and exponential. Then we tabulate the various values of β we get for the different values of risk aversion and different utility functions. This is what the graphs and analysis tell us about the statistical model selection criteria for choosing ARCH. Next, we report the numerical results.

[13]For simplicity, we do not allow asymmetric quadratic effects for positive versus negative errors. However it could be an interesting and substantial extension for future work.

[14]The tangent line, $L_i = \beta_0 + \beta_3 \varepsilon_i^2$, guarantees a certain sign of β_1 due to the "steep" curve on the left of the peak(s) coming out of one side of the tangent line itself – here it is lies below its tangent line, otherwise the sign of β_1 is uncertain.

[15]For a symmetric "hump" shape, β_2 is zero in that positive and negative errors of the same magnitude make same contribution to utility.

4.5 Results

We report the empirical results in Table 4.2.

In Panel A of Table 4.2, we find that for the quadratic utility function both β_1 and β_2 have a constant negative sign while β_3 has a constant positive sign. As the risk aversion coefficient γ rises (from 0.1 to 0.9) β_1 and β_2 tend to increase and β_3 tends to decrease with major changes. The coefficient value of β_3 is higher than those of β_1 and β_2 for each γ.

Panel A: Quadratic utility function (with risk aversion parameter γ)					
Reg. coef.	$\gamma = 0.1$	$\gamma = 0.3$	$\gamma = 0.5$	$\gamma = 0.7$	$\gamma = 0.9$
β_0	0.121260	0.057941	0.045220	0.039728	0.036645
	(0.005971)	(0.001969)	(0.001169)	(0.000826)	(0.000635)
β_1	-0.064303	-0.021203	-0.012584	-0.008891	-0.006839
	(0.184454)	(0.060822)	(0.036098)	(0.025504)	(0.019619)
β_2	-0.071649	-0.023626	-0.014022	-0.009907	-0.007621
	(0.090137)	(0.029722)	(0.017640)	(0.012463)	(0.009587)
β_3	0.338404	0.111586	0.066226	0.046789	0.035993
	(0.936059)	(0.308658)	(0.183189)	(0.129424)	(0.099561)
Panel B: Exponential utility function (with risk aversion parameter A)					
Reg. coef.	$A = 0.1$	$A = 0.3$	$A = 0.5$	$A = 0.7$	$A = 0.9$
β_0	-0.989606	-0.984289	-0.979001	-0.973742	-0.968510
	(0.000761)	(0.000757)	(0.000753)	(0.000749)	(0.000745)
β_1	-0.008104	-0.008060	-0.008017	-0.007974	-0.007931
	(0.023516)	(0.023390)	(0.023264)	(0.023139)	(0.023015)
β_2	-0.009569	-0.009518	-0.009467	-0.009416	-0.009365
	(0.011492)	(0.011430)	(0.011369)	(0.011307)	(0.011247)
β_3	0.042437	0.042209	0.041982	0.041757	0.041532
	(0.119340)	(0.118698)	(0.118061)	(0.117426)	(0.116796)

Notes: We repeat the exercises for different values of risk aversion and for the quadratic and exponential utility cases. We tabulate the various $\beta' s$ values for the different risk aversion values and different utility functions. A precision of 6 decimal digits is required due to the very small numerical results. The standard errors in parentheses are placed below the coefficient estimates.

Table 4.2: Regression coefficients for different risk aversion values and for different utility functions

The results in Panel B of Table 4.2 show that, for the exponential utility function, when γ varies from 0.1 to 0.9 both β_1 and β_2 which are negative become higher and β_3 which is positive becomes lower with minor changes. β_3 is still the one with the highest coefficient values, higher than those of β_1 and β_2 for all risk aversion coefficients A.

As reported above, the quadratic and exponential utility functions suggest consistent empirical results: (1) $\beta_1 < 0$, $\beta_2 < 0$, and $\beta_3 > 0$; (2) β_1 and β_2 increase while β_3 decreases when the risk aversion coefficient is higher; (3) for high values of risk aversion β_3 is always higher than β_1 and β_2. The empirical results match we "predicted" by analyzing the shape of the curves drawn in Figure4.2 in Section 4.3.3. Note that the coefficients of β_i $(i = 1, 2, 3)$ are not only in relation to but also defining and interpreting the shape of the graphs, together not separately, where the risk averse investor will have lower expected utility when the conditional variance is underestimated than when it is overestimated.

We see that for high values of risk aversion we get increasing but relatively lower β_1 and β_2 and decreasing but relatively higher β_3. The results seem robust regardless of whether we use quadratic or exponential utility functions. Hence, if the investor is getting more risk averse he should be using a statistical criterion with more weights on MAE and $MANE$ and less weight on MSE though the weight on MSE is relatively higher than those on MAE and $MANE$ but MSE is becoming less heavily weighted in the criterion that is a weighted average of MAE and MSE.

The chapter's aim requires us to re-do the exercise over and over for different values of the risk aversion parameter and hence get different sets of β_i $(i = 1, 2, 3)$. We emphasize that β_1, the weight on $|e_i|$, rises with risk aversion, which implies that the importance of MAE relative to other standard criteria rises with risk aversion.

Finally, although the model is used is irrelevant to the shape of the graphs, we are interested in knowing the out-of-sample predictive abilities of the three forecasting volatility models (GARCH, EGARCH and CGARCH) in Section 4.2.3 using the objective functions described in Section 4.3 to evaluate their forecast performance. Elder and Gannon (1998) claim that the volatility models' ranking based on daily trades are quite different from those based on positions held until expiration in an economic value framework. On the other hand, Diebold (2001) emphasizes that "ranking of forecast accuracy may, of course, be very different across different loss functions and different horizons". Christoffersen and Jacobs (2004) illustrate consistency in the choice of loss functions that is crucial where the loss function used in parameter estimation and model evaluation should be the same and the estimation loss function should be identical across models when comparing models. In the chapter, we find (but do not report) that the GARCH model performs best for both quadratic and exponential utility cases for each risk aversion level. The optimal forecast error criterion the GARCH model suggests is consistent with the optimal forecast error criterion when we pool all observations together to get maximum efficiency for the regression for interests.

4.6 Conclusion

In the chapter, the problem of forecast error under asymmetric utility-based economic loss functions is examined. An optimality criterion for such utility maximization is suggested. The optimal forecast error criterion is analytically and numerically established with the mapping from errors to wealth. Functions of forecast errors roughly equate with maximizing utility under the trading rule. The optimal forecast error criterion lead us to select a volatility model which is very close to that we get from using economic loss function with similar loss. The optimal forecast error criterion provides a simple statistical rule to make economic and financial decisions under uncertainty.

Conditional heterscedasticity in financial markets has been acknowledged in the fact that the financial asset return has thicker tails than a normally distributed variable as defined in ARCH class models. The underlying principle in the ARCH class of models is that they explain the random variation in conditional variance and thus reserve this tail thickening. According to these, we provide a composite assessment criterion of the ARCH forecasting performance, with respect to economic as well as statistical evaluation aspects. Our motivation stems from two points: the first is Christodoulakis and Satchell's (1998) conclusion of poor-out-of-sample ARCH forecasting performance when judged on the basis of traditional forecast accuracy criteria versus its good performance when more advanced procedures such as utility-based criteria are employed; and the second is the fact that utility-based economic criteria are not practical although using these utility-based criteria may well lead to high profits in trading, because no econometrician or statistician would recognize them or be able to analyze their statistical properties in practice.

We suggest some potentially useful areas for further research. An obvious possibility is to see if other asset classes, such as exchange rate, swap, property, commodity and etc., deliver the consistent optimal forecast error criterion. Another is to permit flexible use of a variety of utility and statistical loss functions by allowing for weighted combinations of loss functions, possibly with time varying weights. Finally, it would be very desirable to compare the optimal forecast error criterion in an environment of dynamic rather than static utility maximization.

Chapter 5

Conclusion

Financial market volatility is known to cluster. A volatile period tends to persist for some time before the market returns to normality. The ARCH model proposed by Engle (1982) was designed to capture volatility persistence in inflation. The ARCH model was later found to fit many financial time series and its widespread impact on finance has led to the Nobel Committee's recognition of Robert Engle's work in 2003. The ARCH effect has been shown to lead to high kurtosis which fits in well with the empirically observed tail thickness of many asset return distributions. The leverage effect, a phenomenon related to high volatility brought on by negative return, is often modelled with a sign-based return variable in the conditional volatility equation.

At the time of writing this thesis, the number of conditional volatility studies is considerable and still rising – nearly all empirical work in finance published this decade is involved with conditional volatility in returns. Previous studies of the statistical properties of financial asset price changes have established, as stylized facts, that relative to the normal such distributions display fat details and asymmetry; and that they display evidence of volatility clustering, with periods of large changes being followed by further large changes but then giving way to periods of intervening tranquillity.

This phenomenon, reflecting the persistence of volatility, is related to the heavy tails of marginal and conditional distributions of returns. The ARCH models are useful and easy to implement in this pursuit because it is estimated on the basis of return distribution. The group of ARCH models is a scientific breakthrough and has triggered intense research in the domain of financial econometrics. The important property of ARCH models is their ability to capture the tendency for volatility clustering in financial data.

This thesis has concentrated on three important questions about conditional volatility: if volatility is forecastable, which econometric method will provide the best forecasts? What economic behaviour is consistent with autoregressive heteroskedastic conditional volatility, if any? What optimal statistical forecast error criterion of conditional volatility forecast would lead to the maximization of economic utility? The first two questions are investigated and discussed in the foreign exchange market, while the last question is studied in the derivative market. Conditional volatility in financial markets is deeply annotated by the new ideas, viewpoints, methodologies and theoretical underpinning employed in this thesis, which meets the wide need for better controlling and modelling asymmetric and clustering volatility requested by both researchers and practitioners.

5.1 A summary of the main results

To consider these questions mentioned above, a number of interesting and substantial works in three main chapters composing the thesis have been studied and discussed.

Chapter 2 investigates the out-of-sample predictive ability of 73 competing time series models for the volatility of foreign exchange changes. Using the evaluation criteria of forecast accuracy and efficiency tests, we compare the out-of-sample forecasting performance of the monthly volatility of the US Dollar versus UK Sterling exchange rate from the post-Bretton Woods era to the present day. The empirical results support the stylized facts of volatility. Historical volatility models are superior to ARCH class models. However, ARCH class models take predominance where over-predictions are more heavily penalized. The various model ranks are shown to be sensitive to the error statistics used to assess the accuracy of the forecasts. The frequency of the data, the type of information used, the forecast horizon, the forecast model and the evaluation criterion are all important variables in volatility forecasting. There is no single forecasting model suitable for all purposes.

Chapter 3 presents a theoretical underpinning to the well established empirical stylized fact that asset returns in general and the spot foreign exchange returns in particular display predictable volatility characteristics. Adopting Moore and Roche's habit persistence version of the Lucas model we find that both the innovation in the spot foreign exchange returns and the foreign exchange returns itself follow "ARCH" style processes. Using the impulse response functions we show that the baseline simulated foreign exchange series has "ARCH" properties in the quarterly frequency that match well the "ARCH" properties of the empirical monthly estimations in Chapter 2, in that when we scale the x-axis to synchronize the monthly and quarterly responses we find similar impulse responses to one unit shock in variance. The impulse response functions for the ARCH processes we estimate "look the same" with an approximately monotonic decreasing fashion. The Lucas two-country monetary model with habit can generate realistic conditional volatility in spot foreign exchange returns.

Chapter 4 proposes an optimal forecast error criterion for utility maximization under an option trading rule. Analysing the quadratic and exponential utility functions, which give the "utility" or "loss" of the cumulated profits from the repeated daily S&P 500 index option trade, we find that both utility cases are asymmetric and peak when the forecast conditional variance equals the actual conditional variance (forecast error is zero). In the sense that the expected utility is a declining function of forecast error, we regress expected utility on forecast error and find that the coefficients in the regression depend on the parameters in the economic problem an investor faces, including the risk aversion parameter and the level of conditional variance. Taking the averaged form of the regression gives the approximate optimal forecast error criterion in terms solely of recognizable statistical loss functions like MAE, MSE etc. We repeat this procedure for different levels of risk aversion and study how the regression coefficients change when the risk aversion parameter changes. The empirical results show that for a more highly risk averse investor the optimal forecast error criterion is a weighted average of MAE and MSE but which weights MSE less heavily. The optimality forecast error criterion based on functions of forecast errors for utility maximization under asymmetric loss provides a simple rule for making economic and financial decisions under uncertainty.

5.2 Suggestions for future research

Volatility plays an important role in investment, portfolio construction, option pricing, hedging, risk management and monetary policy making. Since financial risk and decision-making under uncertainty are commonly assessed in terms of asset volatility, the ability of providing accurate capture and assessment of future risks acquires great importance.

The areas of future work concerning three issues about conditional volatility this thesis has studied, which were suggested previously in detail in each conclusion part of Chapters 2, 3, and 4, are revisited and extended briefly as follows:

Financial market volatility is clearly forecastable. The development of different forecasting models for volatility is guided by the stylized facts observed in the data.

This leads to a large array of alternative models available to practitioners, including all models employed in Chapter 2. However, alternative models should be considered as complements for each other rather than competitors. It will be useful to compare and contrast the full range of time series models with options-based volatility forecasts in terms of out-of-sample forecasting. Inspection of the data and testing for stylized facts naturally appear to be important first steps for practitioners in order to determine which model is best suited for any given situation. It will be also useful to apply the models to other financial and commodities returns.

Conditional volatility, as acknowledged, is one of the most prominent features of volatility in financial markets. The economic factors behind the observed volatility clustering in returns helps understand and improve modelling and forecasting accuracy. A potentially useful area for future research is whether modelling and forecasting power can be enhanced by using exogenous variables. For example, both Chapter 3 and Campbell and Cochrane (1999) provided a habit persistence explanation to predictable conditional volatility. Bittlingmayer (1998) linked volatility to macroeconomic news and systemwide factors; Spiro (1990) and Glosten et al. found a positive relationship between interest rates and volatility; Bollerslev and Jubinski (1999) found a positive relationship between trading volume and volatility; Hamilton and Lin (1996) showed that volatility is higher during recessions. Taylor and Xu (1997) fit 120 seasonal factors to the conditional variance. What the literature has not yet shown is how these relationships improve conditional volatility modelling and forecasting. Hence, econometric techniques are needed in capturing and controlling for heteroscedastic and clustered volatility in financial markets.

As far as is known, models that perform poorly on statistical grounds may still yield a profit if used for trading, and vice versa. A number of volatility forecasting studies have led to the perception that the ARCH-type models provide poor out-of-sample forecasts of volatility when judged on the basis of traditional forecast accuracy criteria versus its good performance when more advanced procedures such as expected utility or profit maximization are employed (see Knight and Satchell, 1998). The optimality forecast error criterion based on functions of forecast errors for utility maximization proposed in Chapter 4 provides a simple rule for making economic and financial decisions under uncertainty. In further work, an obvious possibility is to see if other asset classes, such as exchange rate, swap, property, commodity and etc., deliver the consistent optimal forecast error criterion as in Chapter 4. Another is to permit flexible use of a variety of utility and statistical loss functions by allowing for weighted combinations of loss functions, possibly with time varying weights. Finally, it would be very desirable to compare the optimal forecast error criterion in an environment of dynamic rather than static utility maximization.

There are many old issues that have been around for a long time and many new adventures that are currently underway as well (see Poon and Granger, 2003). Given the importance of the issues addressed in models that allow for structural breaks, in future research it would be useful to focus on volatility models and develop new techniques that will make them more readily available to practitioners (see Knight and Satchell, 1998). For example, there are more tests on the use of absolute returns models in comparison with squared returns models in modelling and forecasting volatility; a multivariate approach to volatility modelling and forecasting where cross correlation and volatility spillover may be accommodated; the realized volatility approach noticeably driven by Andersen, Bollerslev, Diebold, and various co-authors estimating and forecasting volatility risk premium where both spot and option price data are used simultaneously (e.g., Chernov and Ghysels 2000), and the Bayesian and other methods used to estimate stochastic volatility models (e.g. Jones 2001), etc.

Appendix
Appendix A

Author (s)	FX	Data Period	Data Freq	Forecasting Models	Forecasting Horizon	Forecasting Evaluation	Main Conclusions & Comments
Alizadeh, Brandt & Diebold (1999)	GBP/USD DEM/USD SFR/USD JPY/USD CAD/USD (Futures)	1/1/1978-31/12/1998	D	Stochastic Volatility Models (One-Factor and Two-Factor)	One month ahead.	Asymptotic standard errors (ASE), QML (Log Abs. Return) & QML (Log Range) & Exact ML estimators	The Range-based Gaussian quasi-maximum likelihood estimation of stochastic volatility models provides the best of both worlds: simplicity and efficiency. Two-factor stochastic volatility model is desirable.
Amin & Ng (1997)—for interest rate not FX	3M Eurodollar (Futures & futures options)	1/1/88 – 1/11/92	D	Implied $_{AmericanAllCall+Put}$ (WLS, 5 variants of the HJM model) HIS	20 days ahead	R^2 is 21% for implied and 24% for combined. $H_0 : \alpha_{implied} = 0$, $\beta_{implied} = 1$ cannot be rejected with robust SE.	Interest rate models that incorporate volatility term structure perform best. Interaction term capturing rate level and volatility contribute additional forecasting power.
Andersen & Bollerslev (1998)	DEM/USD JPY/USD	In: 1/10/87–30/9/92 Out: 1/10/92-30/9/93	D (5 min)	GARCH (1,1)	1 day ahead, use 5-min returns to construct "actual vol"	R^2 is 5 to 10% for daily squared returns, 50% for 5-min square returns.	GARCH process tracks volatility far better when ex-post volatility is measured on an intra-daily basis than on a daily basis. R^2 increases monotonically with sample frequency.
Andersen, Bollerslev & Diebold (2003)	DEM/USD, S&P500 index; 30-year U.S. Treasury bond yield	12/86-6/99 1/90-12/02 1/90-12/02	D (5-min) Tick-by-tick (5-min) Tick-by-tick (5-min)	HAR Model HAR-RV-CJ Models HAR-RV-J Models	At the daily, weekly, monthly and quarterly forecast horizons	R^2	The jump component of the price process appears to be distinctly less persistent than the continuous sample path component. Explicitly including the jump measure as an additional explanatory variable in an easy-to-implement reduced form model for realized volatility results in highly significant jump coefficient estimates at the daily, weekly and quarterly forecast horizons. When the non-parametric continuous sample path and jump variability measures are included individually in the same forecasting model, only the former measures carry any predictive power for the future realized volatilities.

Andersen, Bollerslev, Diebold & Labys (2000)	DEM/USD JPY/USD (Spot Exchange Rate)	1/12/1986-30/11/1996	D (5 min)	GARCH Long-memory Model	h=1, 5, 10, 15 and 20 days horizon	R^2; Realized Standard Deviations (stdd and stdy); Realized Logarithmic Standard Deviations (lstdd and lstdy) and Realized Correlations (corr)	A multivariate linear Gaussian long memory model is appropriate for daily realized logarithmic standard deviations and correlations. Such a model could result in important improvements in the accuracy of volatility and correlation forecasts and related value-at-risk type calculations.
Andersen & Bollerslev & Diebold & Labys (2001)	JPY/USD DEM/USD JPY/DEM (Spot Exchange Rate) (Reuters FXFX Quotes)	1/12/1986-30/6/1999 In: 1/12/86-1/12/96, 10 years Out: 2/12/1996-30/6/1999, 2.5 years	D (30min)	GARCH (1, 1), VAR-RV, VAR-ABS, RiskMetrics, VAR-RV+RiskMetrics, VAR-RV+GARCH, VAR-RV+VAR-ABS (VAR-RV: forecasts from a long-memory vector autoregression for daily realized volatility; VARABS: forecasts from a long-memory vector autoregression for daily log absolute returns; RiskMetric: forecasts from an exponential smoothing model applied to squared daily returns.)	1 & 10-day-ahead. "Actual vol" derived from 30-min returns.	ASE; SE; R^2	Most procedures for modelling and forecasting financial asset return volatilities, correlations and distributions rely on restrictive and complicated parametric multivariate ARCH or stochastic volatility models, which often perform poorly at intraday frequencies. A long-memory Gaussian VAR for a set of daily logarithmic realized volatilities produces strikingly successful volatility forecasts, which surpass those from conventional GARCH and related approaches. It also generates well-calibrated density forecasts and associated quantile or VaR, estimates.
Andersen & Bollerslev & Diebold & Labys (2002)	¥/US$, DM/US$ (Reuters FXFX Quotes)	1/12/1986-30/6/1999 In: 1/12/86-1/12/96, 10 years Out: 2/12/1996-30/6/1999, 2.5 years	Tick (30min)	VAR-RV, AR-RV, FIEGARCH-RV GARCH-D, RM-D, FIEGARCH-D VAR-ABS	1 and 10 days ahead. "Actual vol" derived from 30-min returns.	1-day ahead R^2 ranges between 27-40% (1-day ahead); and 20-33% (10-day ahead).	RV is realized volatility, D is daily return, and ABS is daily absolute return. VAR allows all series to share the same fractional integrated order and cross series linkages. Forecast improvement is largely due to the use of high frequency data (and realized volatility) instead of the models.
Andersen, Bollerslev & Lange (1999)	DEM/USD (Reuters Quotes)	1/12/1986 – 30/11/1996 In: 1/10/87-30/9/92	5 min	GARCH (1,1) at 5-min, 10-min, 1-hr, 8-hr, 1-day, 5-day, 20-day interval.	1, 5 and 20 days ahead, use 5-min returns to construct "actual vol"	RMSE, MAE, HRMSE, HMAE, LL	Standard volatility models generally provide good forecasts of this economically relevant volatility measure. Moreover, the use of high frequency returns significantly improves the longer run inter-daily volatility forecasts, both in theory and practice. High frequency returns and high frequency GARCH (1,1) models

							improve forecast accuracy. But, for sampling frequencies shorter than 1 hour, the theoretical results and forecast improvement break down.
Baillie, Cecen, Erkal & Han (2004)	GBP/USD DEM/USD SFR/USD	2/1/1986-15/7/1986 (Spot exchange rates from MMS.) 1/1/1996-1/1/1997 (Reuter FXFX quotes from Olson & Associates of Zurich)	H 30-min	MA-FIGARCH	At 30-min interval	SE; Ljung–Box test statistics; BDSL Statistics	High-frequency currency returns volatility is well represented by a FIGARCH model. The estimates of the long memory parameter are remarkably consistent across time aggregations and currencies and are suggestive of self-similarity. At the same time there is some evidence of a small remaining amount of non-linear temporal dependence and it is found to be too weak to be exploitable for forecasting purposes. A relatively simple MA(1)-FIGARCH(1, *d*, 1) model represents the high-frequency currency market data surprisingly well; with the estimates of the long memory volatility parameter being quite close to non-parametric estimates obtained in previous studies.
Balaban (2004)	DEM/USD	2/1/1974-30/12/1997	D	ARCH (p) GARCH (1, 1) EGARCH (1, 1) GJR-GARCH (1, 1)	Multi-step ahead forecasts based on a 72-month rolling estimation procedure	1. Symmetric: ME, MAE, MSE, MAPE; 2. Asymmetric: MME (U), MME (O) and LE; 3. Regression of Unbiasedness: R^2, and SE	The standard symmetric GARCH (1,1) model provides relatively good forecasts of monthly exchange rate volatility whereas the asymmetric GJR-GARCH (1,1) model seems to be a poor alternative.
Barucci & Reno (2002) "On Measuring Volatility and the GARCH Forecasting Performance"	DEM/USD JPY/USD	1/10/1992-30/9/1993	Tick-by-tick quotes	The method is based on Fourier analysis and therefore on the integration of the time series rather than on its differentiation; GARCH (1, 1)	One-step-ahead daily forecasts	RMSE, HRMSE, Fourier estimator & R^2	The method based on Fourier analysis is to compute volatility; the main feature of this method is that it is based upon integration instead of differentiation of the time series, so that it naturally exploits the time structure of high frequency data by including all the observations in the volatility computation. Using simulated time series, this method performs better than the cumulative squared intraday returns in measuring integrated volatility and, according to

							it, the forecasting performance of the GARCH model is improved. When employing the Fourier method, GARCH forecasts turn out to be more accurate than those associated with the sum of squared intraday returns. Moreover the forecasting properties of the GARCH model are evaluated to be better if the Fourier estimator is employed, instead of the cumulative squared intraday returns, to measure integrated volatility.
Bauwens & Omrane & Giot (2003)	EURO/USD	15/5/2001-14/11/2001	5-minute Tick-by-tick	EGARCH	Every 5-minute interval by dividing each day into 288 five-minute intervals	P-value; Wald statistics; Ljung-Box statistics	EGARCH study of the sensitivity of volatility with respect to the news announcements and market activity. Volatility increases in the pre-announcement periods, particularly before scheduled events; the reaction of volatility in the post-announcement period is in most cases muted.
Bazdresch & Werner (2005)	Mexican Peso/USD	1/1996-12/2001	M	AR (1); Market (forwards); Random walk; Switching Regime model; Regime switching conditional on low volatility state; Regime Switching duration dependent model; Regime switching variable probability model with multiple variables	1-month ahead.	MSE; SE; Diebold and Mariano (DM) and sign test (ST) tests; p-value	The forecasting ability of the regime-switching model is no better than that of a random walk with drift. The best predictor is the market. The AR (1) model is just about as good as the random walk. The variable probability model is the worst predictor. The switching regime model restricted to staying in the low volatility regime is the second best predictor, behind the forwards. Duration dependence or fundamentally driven transition probabilities do not improve the model's forecasting power.
Benavides (2004)	Mexican peso/USD Futures, options and spot prices.	Futures and spot rate: 3/9/2001-5/1/2004 Options: 2/1/2002-5/1/2004	D D	A univariate GARCH: GARCH (1, 1); a multivariate ARCH (the BEKK model): Bi-variate BEKK (1, 1), Tri-variate BEKK (1, 1); two option implied volatility models: BS and BAW option implied models; a	1-day ahead at daily interval.	MSE; DW test and R^2	The option implied models are superior to the historical models in terms of accuracy and the composite forecast model was the most accurate one (compared to the alternative models) having the lowest MSE. It is recommended that a composite forecast model is used if both types of data are available i.e. the time-series

		In: 2/1/2002-30/12/2002 Out: 31/12/2002-5/1/2004		composite forecast model			(historical) and the option implied.
Bera & Higgins (1997)	Daily SP500, Weekly $/£, Monthly US Ind Prod	SP 1/1/88-28/5/93 $/£ 12/12/85-28/2/91 IndProd 1/60–3/93	D W M	GARCH Bilinear model	One step ahead. Reserve 90% of data for estimation	Cox MLE RMSE (LE: logarithmic error)	Considers whether heteroskedasticity is due to bi-linearity in the level. Forecasting results show strong preference for GARCH.
Brooks(1997) Brooks & Burke (1998)	DEM/USD JPY/USD CAD/USD	21/3/1973-20/9/1989 In: the first half of the data, 432 observations; Out: the remainders	W	AR(p)-GARCH(r, m): AR (0)-GARCH (0, 0) up to AR (5)-GARCH (5, 5), GARCH (1, 1)	1, 12, and 24 step ahead forecasts at the rolling fixed window of length 432	MSE, MAE	The criteria lead to models which generally provide more accurate forecasts on mean absolute error grounds at short forecasting horizons than a fixed GARCH (1, 1) model, although the GARCH (1, 1) model is still preferable if the forecasts are evaluated using mean squared error.
Calvet & Fisher (2001)	Theoretical Paper (N/A)			Poisson Multifractal Model (PMM)	Theoretical Paper N/A	A grid-free multifractal measure in continuous time	(1) The paper has developed analytical forecasting methods for a new stochastic process, the Poisson multifractal model (PMM). The PMM provides a fully stationary version of the compound process. It parsimoniously captures the volatility persistence, moment-scaling and thick tails that characterize many financial time series. (2) The PMM also supports differing beliefs regarding the long-run behaviour of volatility. (3) The multifractal approach can generate tractable models that match important features of the data, and offers promising new directions for future research in econometrics and financial theory.
Calvet, Fisher & Thompson (2005)	DEM/USD JPY/USD GBP/USD	1/6/73-30/10/03 Out: 1990-2003 (The Deutsche Mark is replaced by the Euro at the beginning of	D	Markov-Switching Multifractal (MSM); CC-GARCH	Two-step estimation at 1 and 5-day horizon	BIC; BIC p-value; CVM statistic; SE	Bivariate MSM performs well in- and out-of-sample relative to a standard benchmark, CC-GARCH. In contrast to CC-GARCH, bivariate MSM captures well the conditional distribution of a variety of currency portfolios. MSM also provides reasonable measures of value-at-risk (VaR), while CC-GARCH tends to

		1999)					underestimate the riskiness of a currency position.
Campa & Chang (1998)	USD/DEM USD/JYP DEM/JYP Option	3/1/1989- 23/5/1995	D	His Implied RiskMetrics-EWMA GARCH (1, 1)	1 and 3 months ahead	RMSE; WRMSE; R^2; Forecast Bias; Wald test	Implied correlation outperforms these alternative forecasts. In combinations, implied correlation always incrementally improves the performance of other forecasts, but the converse is not the case; in certain cases, historically based forecasts contribute no incremental information to implied forecasts. The superiority of the implied correlation forecast holds even when forecast errors are weighted by realized variances.
Cecen & Erkal (1996) (USA)	GBP/USD DEM/USD SFR/USD YEN/USD (Spot rate)	0.00 a.m. 2/1/1986-- 11: 00 a.m. 15/7/1986	H	Non-linear stochastic models (ARCH or GARCH)	Fixed time intervals at the rolling window of 3191 trading hours	BDS test	Non-linear stochastic models can be used for conditional volatility forecasts.
Chen & Leung (1998)	NT/USD	16/12/1994- 8/3/1995	5 min Intraday	EGARCH-M model	1-hour ahead.	R^2; GED	EGARCH specification of the conditional volatility equation appears to provide a good characterization of the stochastic behaviour of the intraday volatility. The significant EGARCH process in the conditional volatility equation indicates that future intraday volatility is also predictable using past information. The EGARCH-M model combined with the GED appears to provide a good characterization of the stochastic behaviour of the intraday exchange rate data and can be used to predict future intraday movements and volatility patterns.
Christoffersen & Diebold (1997)	4 Dollar Exchange Rates: DEM/USD GBP/USD JPY/USD FRF/USD 4 Equity Indexes;	1/1/73-1/5/97	D	Model-Free Assessment of Volatility Forecastability	Non-overlapping h-day returns, h=1, 2, 3, …, 20	Runs Tests for testing independence of the hit sequence; Markov Transition Matrix Eigenvalues and First-Order Correlations for persistence measures	Volatility forecastability varies not only with the horizon, but also with the model. To address this problem, the paper develops a model-free procedure for measuring volatility forecastability across horizons. Volatility forecastability decays quickly with horizon. Equity & FX Markets: For aggregation

	U.S. 10 year Treasury Bond						levels of less than ten days equity return volatility is significant forecastable, and conversely for aggregation levels greater then ten days. Forecastibility decrease rapidly from 1 to 10 days. Bond Markets: volatility is more forecastable in bond markets than elsewhere. Substantially more volatility forecastability than in the equity or foreign exchange markets, with some forecastability out as far as 15-20 trading days.
Christoffersen & Diebold (1999)	4 exchange rates: DEM/USD GBP/USD JPY/USD FRF/USD 4 stock indexes U.S. 10-year Treasury	1/1/73-1/5/97	D	Model-Free: Assessment of Volatility Forecastability by assessing the adequacy of interval forecasts	h = 1, 2, 3, ..., 20 days through non-overlapping h-day returns.	Runs Tests for assessing independence of the hit sequence & Markov Transition Matrix Eigenvalues for measuring volatility forecastability.	Volatility forecastability decays quickly with horizon. If the horizon of interest is more than ten or twenty days, depending on the asset class, then volatility forecasts may not be of much importance. (Equity and Foreign Exchange Markets: for horizons of less than ten trading days, equity return volatility is significantly forecastable, and conversely for horizons greater than ten days; Bond Markets: volatility is more forecastable in bond markets than elsewhere. Substantially more volatility forecastability than in the equity or foreign exchange markets, with some forecastability as far ahead as, say, 15-20 trading days.)
Christoffersen and Diebold (2000)	4 stk indices 4 ex rates US 10 year T-Bond	1/1/73 – 1/5/97	D	No model. (No rank; Evaluate volatility forecastibility (or persistence) by checking interval forecasts.)	1 to 20 days	Run tests and Markov transition matrix eigenvalues (which is basically 1st-order serial coefficient of the hit sequence in the run test).	Equity & FX: forecastibility decreases rapidly from 1 to 10 days. Bond: may extend as long as 15 to 20 days. Estimate bond returns from bond yields by assuming coupon equal to yield.
Cumby, Figlewski and Hasbrouck (1993)	¥/$, Stocks (¥, $), Bonds (¥, $)	7/77 - 9/90	W	EGARCH HIS	1 week ahead, estimation period ranges from 299 to 689 weeks.	R^2 varies from 0.3% to 10.6%.	EGARCH is better than naive models in forecasting volatility though R-square is low. Forecasting correlation is less successful.

DeGennaro & Shrieves (1995)	JPY/USD (Screen quote)	1/10/92-30/9/93	H	GARCH	At hourly intervals (day-of-week & hour-of-day)	SE	A GARCH model is used to test a hypothesis related to the effects of news arrival on conditional return variance processes and to capture the effects of specified news events on return volatility, while controlling for known seasonalities in conditional variance and for information effects implicit in trading activity levels. The foreign exchange rate volatility increases in the hour prior to news arrival, consistent with the presence of a relatively higher rate of private information arrival just prior to impending news announcements.
DeGennaro & Shrieves (1997)	JPY/USD	1/10/92-30/9/93	D (10-min)	GARCH	10-minute interval of the day	Measurement error	GARCH provides a natural approach to testing hypotheses related to the effects of information arrival on mean return and conditional return variance processes. Both private information and news effects are important determinants of exchange rate volatility. Unexpected quote arrival positively impacts foreign exchange rate volatility, which is consistent with the interpretation that unexpected quote arrival serves as a measure of informed trading.
Dunis, Laws and Chauvin (2000)	DM/¥, £/DM, £/$, $/CHF, $/DM, $/¥	In: 2/1/91-27/2/98 Out: 2/3/98-31/12/98	D	GARCH (1,1) AR (10)-Sq returns AR (10)-Abs returns SV (1) in log form HIS 21 or 63 trading days 1- & 3-M forward Implied $_{ATMquotes}$ Combine Combine (except SV)	1 and 3 months (21 & 63 trading days) with rolling estimation. Actual volatility is calculated as the average absolute return over the forecast horizon.	RMSE, MAE, MAPE, Theil-U, CDC (Correct Directional Change index)	No single model dominates though SV is consistently worst, and implied always improves forecast accuracy. Recommend equal weight combined forecast excluding SV.

Ederington & Guan (2000) "Forecasting Volatility"	DEM/USD	1/1/71-30/6/97	D	GW MAD GWSTD, GARCH, EGARCH AGARCH HISMAD, n, HISSTD, n	n =10, 20, 40, 80 & 120 days ahead estimated from a 1260-day rolling window; parameters re-estimated every 40 days. Use daily squared deviation to proxy "actual" vol.	RMSE, MAE	Volatility aggregated over a longer period produces a better forecast. Absolute returns models generally perform better than square returns models (except GARCH>AGARCH). As horizon lengthens, no procedure dominates. GARCH & EGARCH estimations were unstable at times.
Ederington & Guan (2004) "Forecasting Volatility"	DEM/USD S&P 500 Index 3-month Eurodollar rate 10-year Treasury Bond rate Five equities	1/1/71-6/30/97 7/2/62-12/29/95 1/1/73-6/20/97 1/2/62-6/13/97 7/2/62-12/30/94	D D D D D	LSD models: STD, Riskmetric's EWMA, GARCH (1,1); AGARCH; EGARCH; Two regression models: (A-) RLS and (A-) GEN models.	n =10, 20, 40, 80 & 120 days ahead estimated from a 1260-day rolling window; parameters re-estimated every 40 days. Use daily squared deviation to proxy "actual" vol.	RMSFE; MAFE	(1) No one model dominates at all horizons in all nine markets but one certainly stands out from the others. (2) Models based on absolute return deviations generally forecast volatility better than otherwise equivalent models based on squared return deviations - though not for GARCH models. (3) GARCH (1,1) generally yields better forecasts than the historical standard deviation and exponentially weighted moving average models but between GARCH and EGARCH there is no clear favourite.
Edey and Elliot (1992)	Futures options on A$ 90d-Bill, 10yr bond, Stock index A$/US$ (Options)	Futures options: inception to 12/88 A$/US$ option: 12/84 – 12/87	W W	Implied $_{BK\ NTM,\ call}$ Implied $_{BK\ NTM,\ put}$ (No rank, 1 call and 1 put, selected based on highest trading volume)	Option maturity up to 3M. Use sum of (return square plus Implied t+1) as "actual vol" Constant 1M. Use sum of weekly squared returns to proxy "actual vol".	Regression (see comment). In most cases a implied>0 and b implied<1 with robust SE. For stock index option b implied=1 cannot be rejected using robust SE.	R^2 cannot be compared with other studies because of the way "actual" is derived and lagged squares returns were added to the RHS.

Ederington & Guan (2000) "Forecasting Volatility"	DEM/USD	1/1/71-30/6/97	D	GW MAD GWSTD, GARCH, EGARCH AGARCH HISMAD, n, HISSTD, n	n =10, 20, 40, 80 & 120 days ahead estimated from a 1260-day rolling window; parameters re-estimated every 40 days. Use daily squared deviation to proxy "actual" vol.	RMSE, MAE	Volatility aggregated over a longer period produces a better forecast. Absolute returns models generally perform better than square returns models (except GARCH>AGARCH). As horizon lengthens, no procedure dominates. GARCH & EGARCH estimations were unstable at times.
Ederington & Guan (2004) "Forecasting Volatility"	DEM/USD S&P 500 Index 3-month Eurodollar rate 10-year Treasury Bond rate Five equities	1/1/71-6/30/97 7/2/62-12/29/95 1/1/73-6/20/97 1/2/62-6/13/97 7/2/62-12/30/94	D D D D D	LSD models: STD, Riskmetric's EWMA, GARCH (1,1); AGARCH; EGARCH; Two regression models: (A-) RLS and (A-) GEN models.	n =10, 20, 40, 80 & 120 days ahead estimated from a 1260-day rolling window; parameters re-estimated every 40 days. Use daily squared deviation to proxy "actual" vol.	RMSFE; MAFE	(1) No one model dominates at all horizons in all nine markets but one certainly stands out from the others. (2) Models based on absolute return deviations generally forecast volatility better than otherwise equivalent models based on squared return deviations - though not for GARCH models. (3) GARCH (1,1) generally yields better forecasts than the historical standard deviation and exponentially weighted moving average models but between GARCH and EGARCH there is no clear favourite.
Edey and Elliot (1992)	Futures options on A$ 90d-Bill, 10yr bond, Stock index A$/US$ (Options)	Futures options: inception to 12/88 A$/US$ option: 12/84 – 12/87	W W	Implied $_{BK\ NTM,\ call}$ Implied $_{BK\ NTM,\ put}$ (No rank, 1 call and 1 put, selected based on highest trading volume)	Option maturity up to 3M. Use sum of (return square plus Implied t+1) as "actual vol" Constant 1M. Use sum of weekly squared returns to proxy "actual vol".	Regression (see comment). In most cases a implied>0 and b implied<1 with robust SE. For stock index option b implied=1 cannot be rejected using robust SE.	R^2 cannot be compared with other studies because of the way "actual" is derived and lagged squares returns were added to the RHS.

Edmonds & So (2004)	CAD/USD FRF/USD DEM/USD JPY/USD SFR/USD GBP/USD (Spot and forward rates)	1973-1991 1976-1991	W M	ARCH AR	Not mentioned.	Forecast errors; Phillips–Perron tests	Not only is it inappropriate to characterize all price series similarly (as ARCH, or as GARCH, or not), but any testing to investigate such processes may result in erroneous conclusions unless robust tests are used. The exchange rates, like other asset prices, seem far more likely to be characterized by the ARCH (conditional) variance process in a weekly model than in a monthly one.
Engle & Russell (1997)	USD/DEM	1/10/1992-30/9/1993 (Only Tuesdays, Wednesdays and Thursdays are analyzed for the months of May through August.)	H	ACD model	5-14 hour interval in the 4-month sample; One-step ahead	R^2; LM test	When the durations are measured in terms of price change events, the ACD model becomes a volatility model. The ACD model is able to successfully model seasonal time of day effects and stochastic effects. Using a Weibull density for the hazard proved superior to the exponential, WACD model. This model provides a framework in which the instantaneous probability of events can be forecast. (The bid-ask spread should have predictive power for the volatility that is supported by the data.)
Fiess & MacDonald (2002)	DEM/USD JPY/USD USD/GBP	2/1/1986-30/8/1996 21/8/1989-30/8/1996 2/96-8/96	D 15 min	An unrestricted vector autoregressive model (VAR)	1-step ahead	High-Low (HiLo), Close-to-Close volatility (CLCL), Garman-Klass (GK) estimators and GARCH (1, 1) forecasts	A volatility measure based on the daily trading range has an information advantage over volatility measures based on absolute daily returns. HiLo is superior volatility forecast for CLCL. GK estimator comes second, followed by the GARCH forecast. The naive forecast based on the previous day's return performs worst. GARCH forecasts provide only a second best strategy when extreme value estimators are considered.

Figlewski (1997)	S&P 500 3M US T-Bill 20Y T-Bond DM/$	1/47-12/95 1/47 – 12/95 1/50 – 7/93 1/71 – 11/95	M	$HIS_{6,12,24,36,48,60m}$ GARCH (1,1) for S&P and bond yield.	6, 12, 24, 36, 48, 60 months. Use daily returns to compute "actual vol".	RMSE	Forecast of volatility of the longest horizon is the most accurate. HIS uses the longest estimation period is the best except for short rate.
	S&P 500 3M US T-Bill 20Y T-Bond DM/$	2/7/62 – 29/12/95 2/1/62 – 29/12/95 2/1/62 – 29/12/95 4/1/71 – 30/11/95	D	GARCH (1,1) $HIS_{1,3,6,12,24,60\ months}$	1, 3, 6, 12, 24 months	RMSE	GARCH is best for S&P but gave the worst performance in all the other markets. In general, as out of sample horizon increases, the in-sample length should also increase.
Figlewski & Green (1999)	S&P 500 US LIBOR 10 yr T-Bond Yield DM/$	4/1/71-31/12/96 Out: From Feb96	D	$His_{3,12,60 months}$ ES	1, 3, 12 months for daily data.	RMSE	ES works best for S&P (1-3 month) and short rate (all three horizons). HIS works best for bond yield, exchange rate and long horizon S&P forecast. The longer the forecast horizon, the longer the estimation period.
		4/1/71-31/12/96 Out: From Jan92	M	$His_{26,60,all months}$ ES	24 & 60 months for monthly data.		For S&P, bond yield and DM/$ it is best to use all available "monthly" data. 5 years worth of data works best for short rate.
Frommel (2004)	FIM/DEM IEP/DEM ITL/DEM PTE/DEM ESP/DEM	1/1/1996-31/12/1998	D	MS-GARCH model; A Single Regime GARCH (Modelling not forecasting)	No mentioned.	The persistence of a volatility shock given by the sum of the coefficients α and β. (The higher $\alpha + \beta$ is, the more time it takes until a shock has died out. It will die out in finite time, if $\alpha + \beta$ is smaller than 1, as soon as it exceeds 1 a volatility shock has permanent impact and unconditional volatility is infinity.)	Regime Switching GARCH models deal with the problem of GARCH that arises if the underlying volatility process is subject to structure breaks, especially shifts in the overall level of volatility because the persistence of volatility shocks is systematically overestimated. The high persistence of volatility shocks in single-regime GARCH models is due to the neglecting of regime changes, which is the model is misspecified. Not only official announcements but also market expectations may affect changes in uncertainty and volatility in the environment of a changing

							exchange rate arrangement.
Frömmel & MacDonald & Menkhoff (2005)	DEM/USD JPY/USD GBP/USD	1/1974-10/2000	M	Markov switching RID model; Pure Markov switching model; Constant coefficient RID model; Random walk with drift	1, 6, 12 months by rolling sample of 10 years	RMSE, MA'E, asymptotic test by Diebold and Mariano (1995)	An MSM approach has been proven much better at describing exchange rates than linear approaches. Structural stability is not good enough to systematically produce out-of-sample forecasting ability. Markov switching models are able to describe the data well, but produce poor forecasts. (Conclusion: although the MSM captures most of the structural instability in the coefficients, there is still some additional source of time variation left. There is a nonlinear relationship between exchange rates and macroeconomic fundamentals.)
Fung, Lie and Moreno (1990)	£/$, C$/$, FFr/$, DM/$, ¥/$ & SrFr/$ (Options on PHLX)	1/84 – 2/87 (Pre crash)	D	$\text{Implied}_{OTM>ATM}$ $\text{Implied}_{vega,elasticity}$ $\text{Implied}_{equalweight}$ HIS_{40days}, Implied_{ITM}	Option maturity; Overlapping periods. Use sample SD of daily returns over option maturity to proxy "actual vol".	RMSE, MAE of Overlapping forecasts.	Each day, 5 options were studied; 1 ATM, 2 just in and 2 just out. Define ATM as S=X, OTM marginally outperformed ATM. Mixed together implied of different contract months.
Fung and Hsieh (1991)	S&P 500, DM/$ US T-bond (Futures and Futures Options)	3/83 – 7/89 (DM/$ futures from 26 Feb 85)	D (15min)	RV-AR (n) Implied BAW NTM Call/Put RV, RW (C-t-C) HL	1 day ahead. Use 15-min data to construct "actual vol".	RMSE and MAE of $\log \sigma$	RV: Realized vol from 15-min returns. AR (n): autoregressive lags of order n. RW (C-t-C): random walk forecast based on close-to-close returns. HL: Parkinson's daily high-low method. Impact of 1987 crash does not appear to be drastic possibly due to taking log. In general, high frequency data improves forecasting power greatly.

Galbraith & Kisinbay (2005)	DEM/USD JPY/USD (Spot rate)	2/1//1987-31/12/1998	Intraday (5-min)	QML-GARCH QML-FIGARCH AR-rv	One-step forecast at 30-day horizon. 5, 10 and 30 min returns used in estimating realized volatility	MSE & p-value	The 5-min realized volatilities provide the strongest results on all data sets. Projections on past-realized variance provide better one-step forecasts than the QML-GARCH and -FIGARCH forecasts, appears to extend to longer horizons up to around 10 to 15 trading days. At longer horizons, there is less to distinguish the forecast methods, but the evidence does suggest positive forecast content at 30 days for various forecast types.
Gavrishchaka and Ganguli (2003)	DEM/USD	1/1/1980-1/1/2000	D	SVM-based models ARCH/HARCH models GARCH (1, 1)	A standard 5-fold cross-validation procedure; a step of 5, 10 and 20 business days; volatilities computed on a weekly interval from daily returns.	R^2; Regularization parameter C, ε-parameter of the loss function, coefficients of the kernel function, and the type of the kernel function itself.	The main advantage of SVM is its ability to handle high-dimensional data so that SVM-based volatility model can model long memory and multi-scale effects without the restrictive assumptions required by other models. SVM can efficiently work with high-dimensional inputs to account for volatility long-memory and multi-scale effects. SVM can perform significantly better or comparable to both naive and GARCH (1,1) models. The advantages of the SVM-based techniques are expected to be much more pronounced in modelling small-scale (intraday) volatilities and high-frequency financial data.
Guo (1996a)	PHLX US$/¥ (Options)	Jan91 – Mar93	D	$\text{Implied}_{\text{Heston}}$ $\text{Implied}_{\text{HW}}$ $\text{Implied}_{\text{BS}}$ GARCH HIS_{60}	Information not available.	Regression with robust SE. No information on R^2 and forecast biasness.	Use mid of bid-ask option price to limit 'bounce' effect. Eliminate 'nonsynchroneity' by using simultaneous exchange rate and option price. HIS and GARCH contain no incremental information. Implied Heston and Implied HW are comparable and marginally better than $\text{Implied}_{\text{BS}}$.

Guo (1996b)	PHLX US$/¥, US$/DM Options (Spot rate)	Jan86 – Feb93	Tick D	Implied$_{HW}$ (WLS, 0.8 <S/X< 1.2, 20<T<60 days) GARCH (1,1) HIS$_{60days}$	60 days ahead. Use sample variance of daily returns to proxy actual volatility.	US$/DM R^2 is 4, 3, 1% for the three methods. (9, 4, 1% for US$/¥. All forecasts are biased $\alpha>0$, $\beta<1$ with robust SE.	Conclusion same as Guo (1996a). Use Barone-Adesi / Whaley approximation for American options. No risk premium for volatility variance risk. GARCH has no incremental information. Visual inspection of figures suggests implied forecasts lagged actual.
Heynen and Kat (1994)	7 stock indices and 5 exchange rates	1/1/80-31/12/92 In: 80-87 Out: 88-92 (87's crash included in in-sample)	D	SV EGARCH GARCH RW	Non-overlapping 5, 10, 15, 20, 25, 50, 75, 100 days horizon with constant update of parameters estimates. Use sample standard deviations of daily returns to proxy "actual vol".	MedSE	SV appears to dominate in index but produces errors that are 10 times larger than (E) GARCH in exchange rate. The impact of 87's crash is unclear. Conclude that volatility model forecasting performance depends on the asset class.
Hu & Tsoukalas (1999)	BEF/DEM LUF/DEM DKK/DEM FRF/DEM GRD/DEM GBP/DEM IEP/DEM ITL/DEM NLG/DEM PTE/DEM ESP/DEM USD/DEM	13/3/1979-30/12/1994: 3-subperiods: 13/3/1979-4/4/1990; 5/4/1990-2/6/1993; 3/6/1993-30/12/1994 (Out)	D	Four Individual Models: GARCH (1,1), EGARCH, IGARCH and MAV; Three combining Models: AVE, OLS and ANN	One-step-ahead combined volatility forecast at a rolling window of 4124 daily observations	RMSE & RMAE	(1). The superior out-of-sample forecasting performance of the EGARCH model is consistent with the nature of the EMS as a managed float regime (2). The ability of the MAV and ANN models to account for the observed volatility during the foreign exchange crisis of August 1993. Probably, models with longer memory are more appropriate to account for the autocorrelation structure generated in crisis periods (3). The forecast combining performance of the ANN model behaves well in the crisis period and gets a good ranking in the post-crisis period. Its performance is much better in terms of absolute prediction errors than in terms of squared prediction errors.

Jorion (1995)	DM/\$, ¥/\$, SrFr/\$ futures (Options on CME)	1/85 – 2/92 7/86 – 2/92 3/85 – 2/92	D	$\text{Implied}_{\text{ATM BS call+put}}$ GARCH (1,1), MA_{20}	1 day ahead & option maturity. Use squared returns and aggregate of square returns to proxy actual volatility.	R^2 is 5% (1-day) or 10-15% (option maturity). With robust SE, $\alpha_{implied} \succ 0$ and $\beta_{implied} \prec 1$ for long horizon and is unbiased for 1-day forecasts.	Implied is superior to the historical methods and least biased. MA and GARCH provide only marginal incremental information.
Jorion (1996)	DM/\$ futures (Options on CME)	Jan85-Feb92	D	$\text{Implied}_{\text{Black,ATM}}$ GARCH (1,1)	1 day ahead, use daily squared to proxy actual volatility.	R^2 about 5%. H0: $\alpha_{implied} = 0$, $\beta_{implied} = 1$ can not be rejected with robust SE.	R^2 increases from 5% to 19% when unexpected trading volume is included. Implied volatility subsumed information in GARCH forecast, expected futures trading volume and bid-ask spread.
Kilian and Taylor (2003)	CAD/USD CHF/USD FRF/USD DEM/USD JEN/USD ITL/USD GBP/USD	1973.I–1998.IV	Q	Exponential smooth Transition autoregressive (ESTAR) model	2 or 3-year ahead at K=1, 4, 8, 12 and 16 quarters horizon.	P values; MSE; SE of the regression; DM test; And bootstrap test	The power of tests of the random walk hypothesis against forecast models should increase at longer forecast horizons, making it easier to detect predictability using long horizon forecasts. With a bootstrap test, given ESTAR real exchange rate dynamics, strong evidence of predictability is found at horizons of 2 to 3 years, but not at shorter horizons.
Klaassen (1998)	US\$/£, US\$/DM and US\$/¥	3/1/78 – 23/7/97 Out: 20/10/87-23/7/97	D	RSGARCH RSARCH GARCH (1,1)	1 and 10 days ahead. Use mean adjusted 1-and 10-day return squares to proxy actual volatility.	MSE of variance, regression though R^2 is not reported.	GARCH (1,1) forecasts are more variable than RS models. RS provides statistically significant improvement in forecasting volatility for US\$/DM but not the other exchange rates.
Lee (1991)	\$/DM \$/£ \$/¥, \$/FFr \$/C\$ (Fed Res Bulletin)	7/3/73-4/10/89 Out: 21/10/81 – 11/10/89.	W (Wed, 12pm)	Kernel (Gaussian, Truncated) Index (combining ARMA and GARCH) EGARCH (1,1) GARCH (1,1) IGARCH with trend	1 week ahead (451 observations in sample and 414 observations out-of-sample)	RMSE, MAE. It is not clear how actual volatility was estimated	Nonlinear models are, in general, better than linear GARCH. The kernel method is best with MAE. But most of the RMSE and MAE are very close. Over 30 kernel models were fitted, but only those with the smallest RMSE and MAE were reported. It is not clear how the non-linear equivalence was constructed. Multi-step forecast results were mentioned but not shown.

Li (2002)	$/DM $/£ $/¥ OTC ATM Options $/£, $/¥ $/DM	3/12/86- 30/12/99 In: 12/8/86- 11/5/95 19/6/94-13/6/99 19/6/94- 30/12/98	Tick (5 min) D D	Implied GK OTC ATM ARFIMA realized (Implied better at shorter horizon and ARFIMA better at long horizon.)	1, 2, 3 and 6 months ahead. Parameters not reestimated. Use 5-min returns to construct "actual vol".	MAE. R^2 ranges 0.3-51% (Implied), 7.3-47%(LM), 16-53% (encompass). For both models, $H_0: \alpha = 0, \beta = 1$ are rejected and typically $\beta < 1$ with robust SE.	Both forecasts have incremental information especially at long horizon. Forcing: a =0, b =1 produces low/negative $R2$ (especially for long horizon). The model realized standard deviation as ARFIMA without log transformation and with no constant, which is awkward as a theoretical model for volatility.
Lopez (2001)	C$/US$ DM/US$ ¥/US$ US$/£	1980-1995 In: 1980-1993 Out: 1994-1995	D	SV-AR (1)-normal GARCH-gev EWMA-normal GARCH-normal, -t EWMA-t AR (10)-Sq, -Abs Constant	1 day ahead and probability forecasts for four "economic events", viz. cdf of specific regions. Use daily squared residuals to proxy volatility. Use empirical distribution to derive cdf.	MSE, MAE, LL, HMSE, GMLE and QPS (quadratic probability scores).	LL is the logarithmic loss function from Pagan and Schwert (1990), HMSE is the heteroskedasticity-adj MSE from Bollerslev and Ghysels (1996) and GMLE is the Gaussian quasi-ML function from Bollerslev, Engle and Nelson (1994). Forecasts from all models are indistinguishable. QPS favours SV-n, GARCH-g and EWMA-n.
Martens (2001)	DEM/USD YEN/USD (Reuters FXFX quotes & Bankers Trust)	1/1/1996— 31/12/1996 1/1/1988-- 31/12/1995	30 min D	Intraday GARCH (1, 1); Monte Carlo simulation	One day ahead at intraday: 2-min and 5 min-intervals.	SE; HMSE; R^2	Modelling intraday returns and volatility improves the out-of-sample forecasting of daily volatility. The higher the frequency of returns used in the GARCH model, the better the out-of-sample daily volatility forecasts are. Only a relatively small sample of intraday returns is sufficient to produce superior daily volatility forecasts from an intraday GARCH (1,1) model at the highest available frequency.

Martens, Chang & Taylor (2002)	DEM/USD JPY/USD (Spot Rates)	1996 In: Jan96-Sept96 Out: Oct96-Dec96	30 min	Two-step approach based on the FFF; GARCH (1, 1); PGARCH	At thirty-minute horizon	RMSE, LL, R^2 and the correlation between realized and forecasted volatility	Explicitly modelling the intraday seasonal pattern improves the out-of-sample forecasting performance. A seasonal estimated from the log of squared returns improves upon the use of simple squared returns, and the flexible Fourier form (FFF) is an efficient way of determining the seasonal. The two-step approach performs only marginally worse than the computationally expensive periodic GARCH model that includes the FFF.
Martens and Zein (2002)	S&P500 Futures, ¥/US$ Futures, Crude oil Futures	Jan 94 - Dec 2000 Jan 96 - Dec 2000 Jun 93 - Dec 2000	Tick	Implied BAW VIX style Log-ARFIMA GARCH	Non-overlapping 1, 5, 10, 20, 30 and 40 days ahead. 500 daily observations in in-sample which expands on each iteration.	Heteroskedasticity adjusted RMSE. R^2 ranges 25-52% (implied), 15-48% (LM) across assets and horizons. Both models provide incremental info to encompassing regr.	Scaled down one large oil price. Log-ARFIMA truncated at lag 100. Based on R2, Implied outperforms GARCH in every case, and beats Log-ARFIMA in ¥/US$ and Crude oil. Implied has larger HRMSE than Log-ARFIMA in most cases. Difficult to comment on the implied's biasness from information presented.
McKenzie (1999)	21 A$ bilateral exchange rates	Various length from 1/1/86 or 4/11/92 to 31/10/95	D	Square vs. power transformation (ARCH models with various lags.)	1-day ahead absolute returns.	RMS, ME, MAE. Regressions suggest all ARCH forecasts are biased. No R^2 was reported.	The optimal power is closer to 1 suggesting squared return is not the best specification in ARCH type models for forecasting purposes.
Moore & Roche (2004) (Northern Ireland & Republic of Ireland)	USD/GBP	1973:1-1998:4	Q	Standard Model; Monetary Model; Habit model; Driftless Random Walk model	4, 8, and 12 quarters ahead at 4-quarter to 12-quarter horizons	RMSE, MSE, SE and p-value	The model produces superior forecasts (in RMSE) relative to those generated from the monetary or random walk models, particularly at longer horizons.

Morana & Beltratti (2004)	DM/USD Yen/USD	1/12/1986-1/12/1996	5 min	Switching regimes models (MS-AR and MS-ARFIMA) and ARFIMA model. MS–AR denotes the two-step Markov switching autoregressive model, ARFIMA denotes the standard ARFIMA model and MS–ARFIMA denotes the two-step Markov switching ARFIMA model.	1-, 5-, and 10-step ahead	R^2, MSE and forecast encompassing criterion; LM, Wald and LR estimators	MS–AR model would seem to be the preferred model. At a longer forecasting horizon, the Markov switching models seems to perform better than the standard ARFIMA model. The forecast encompassing criterion does not allow selection of a clear winner between the MS–AR and the MS–ARFIMA models at longer horizon, although the MS–ARFIMA model shows a better performance than the MS–AR model. A pure long-memory model may be useful for forecasting purposes, even when the true data generating process shows weak dependence. Modelling the break process is not important for very short term forecasting (one-step ahead) once the model considers a long-memory component. However, at longer forecasting horizons (5–10 steps ahead), accounting for both long memory and structural change leads to a superior forecasting model. Neglecting the break process is not important for very short term forecasting once a long memory component is allowed for in the model, but superior forecasts can be obtained at longer horizons by modelling both long memory and structural change.
Neely (1999)	BEF/DEM DKK/DEM FRF/DEM IEP/DEM ITL/DEM NLG/DEM	14/3/1979-31/7/1992	W	Jump-diffusion GARCH model; Absolute value GARCH model	No mentioned	SE; p-values of F statistic and t statistic	Time-varying jump probability and absolute value GARCH models are effective in improving the fit of jump-diffusion models on target zone data. The absolute value GARCH models provide a more robust estimate of the forecast conditional variance. There is some evidence that conditional volatility is higher around the periods of realignments.

Neely (2004)	Futures and Options-on-Futures: DEM/USD JPY/USD CHF/USD GBP/USD; Returns on FX Interest rates	2/1/1987-31/12/1998 In: 1987 to 1991 Out: 1992-1998	D 30-min D	SV, ARIMA, LM-ARIMA, GARCH, and OLS	Overlapping observations at a fixed horizon; One-step-ahead	R^2; SE; Wald test; Bayesian Information Criterion (BIC); Specification error from using the wrong options model and idiosyncratic error from microstructure effects; Tracking Error	Four types of forecasting models are used to examine the informational efficiency of IV. Out-of-sample forecasts from ARIMA, LM-ARIMA, and OLS models of RV are helpful adjuncts to IV. Specifically, IV fails to subsume time series forecasts of RV using either overlapping or fixed horizon estimation. IV's bias as a predictor of RV is still a puzzle, but it does not matter economically.
Poon & Granger (2001) "Forecasting Volatility in Financial Markets: A Review"	A variety of data sets (Different asset classes & markets in different geographical regions)	Various period	Various data frequency / High frequency data	1. Historical (HIS) 2. GARCH 3. Option implied volatility 4. Stochastic volatility	Good practice that the amount of data used to produce volatility forecasts should be at least as long as the forecast horizon	1. Performance measures: ME, MSE, RMSE, MAE, MAPE, MLAE, Theil-U statistic and LINEX 2. Regression based efficiency test (R^2) 3. Making statistical inference	Implied volatility appears to be the most reliable method of forecasting volatility, with GARCH and HIS roughly equal seconds. SV was studied relatively little, presumably because of practical difficulties.
Poon & Granger (2002) "Forecasting Volatility in Financial Markets: A Review"	Different asset classes & markets in different geographical regions	Different period (The choice of this period should be based on the needs of real decision makers but in practice are rarely justified by researchers.)	Various data frequency / High frequency data	Times series volatility forecasting models (1. Predictions based on past standard deviations; 2. ARCH class conditional volatility models; 3. Stochastic volatility models; Options based volatility forecasts	Complication in relation to the choice of forecast horizon is partly due to volatility mean reversion. Volatility forecast accuracy improves as data sampling frequency increases relative to forecast horizon.	1. Measuring forecast errors: ME, MSE, RMSE, MAE, MAPE, MLAE, Theil-U statistic and LINEX; 2. Comparing forecast errors of different models: an asymptotic test, sign test and the Wilcoxon's signed-rank test; 3. Regression based forecast efficiency and orthogonality test (SE, R^2); 4. Using squared return to proxy actual volatility; 5. Further issues in	(1). Four Categories of Volatility Forecasts: HISVOL, GARCH, ISD and SV. (2). ISD provides the best forecasting with HISVOL and GARCH roughly equal, although possibly HISVOL does somewhat better in the comparisons. GARCH dominates ARCH. Models that incorporate volatility asymmetry such as EGARCH and GJR-GARCH perform better than GARCH. But certain specialized specifications, such as fractionally integrated GARCH (FIGARCH) and regime switching GARCH (RSGARCH) do better in some studies. (3). The option implied volatility being a market-based volatility forecast has been shown to contain most information about future volatility. The supremacy among historical time series models depends on the type of asset being modelled.

						forecast evaluation (in-sample & out-of-sample forecasts and HMSE)	But, as a rule of thumb, historical volatility methods work equally well compared with more sophisticated ARCH class and SV models.
Pong, Shackleton, Taylor & Xu (2002)	USD/GBP	In: 7/87-12/93 Out: 1/94-12/98	5-, 30-minute	$\text{Im}\,plied_{ATM,OTCquote}$ (bias adj using rolling regr on last 5 years monthly data) Log-ARMA (2,1) Log-ARFIMA (1,d, 1) GARCH (1,1)	1 month and 3 month ahead at 1-month interval	ME, MSE, regression. R^2 ranges between 22-39% (1-month) and 6-21% (3-month)	Implied, ARMA and ARFIMA have similar performance. GARCH (1, 1) is clearly inferior. Best combination is Implied+ARMA(2,1). Log-AR (FI) MA forecasts adjusted for Jensen inequality. *Difficult to comment on implied's biasness from information presented.*
Pong, Shackleton, Taylor & Xu (2004)	USD/GBP USD/DEM USD/JPY	1/7/1987-31/12/1998 In: 1/7/87-31/12/93 Out: 1/1/94–31/12/98	5 &30-min	FIV A short memory ARMA (2,1); A long memory ARFIMA (1,d, 1); GARCH (1,1)	1 day to 3 months horizons at the intervals equal to the forecast horizons, except for the three-month forecasts	MSE; R^2; Test for incremental information; SE	ARFIMA and ARMA forecasts generally perform better than implied volatilities for short forecast horizons, while implied volatilities produce more accurate forecasts for longer forecast horizons. The GARCH forecasts are the least accurate for most of the evaluations, although incremental information is found to exist over implied volatilities for the shorter forecast horizons.
Radalj & McAleer (2005)	Five Currency Markets (AUD/USD CAD/USD GBP/USD JPY/USD SFR/USD)	6/10/92-15/10/02	W	ARMA(1,1)-GARCH(1,1) GARCH (incorporating speculators into such models);	No information	SE; LM test; ADF test; p values	As the time-varying conditional volatility GARCH model and its variants have been criticized for lacking economic content, incorporating speculators into such models contributes to an accommodation of this criticism. An ARMA (1,1)-GARCH (1,1) model with exogenous variables is used for the conditional volatility equation, which overcomes a severe limitation that they do not accommodate the direction of speculative trades. The empirical evidence supports the hypothesis that speculators have no impact on volatility.
	Two Commodity Markets (oil and gold)	6/10/92-15/10/02	W				
	Two Stock Market Indices (S&P Composite 500 index and the Nikkei 225 index)	6/10/92-15/10/02	W				

Scott and Tucker (1989)	DM/\$, £/\$, C\$/\$, ¥/\$ & SrFr/\$ American (Options on PHLX)	14/3/83-13/3/87 (Pre crash)	Daily Closing Tick	Implied GK (vega, Inferred ATM, NTM) Implied $_{CEV}$	Non-overlapping option maturity: 3, 6 and 9 months. Use sample SD of daily returns to proxy "actual vol".	MSE, R^2 ranges from 42 to 49%. In all cases, α >0, β <1. HIS has no incremental info content.	Simple B-S forecasts just as well as sophisticated CEV model. Claimed omission of early exercise is not important. Weighting scheme does not matter. Forecasts for different currencies were mixed together.
Szakmary, Ors, Kim and Davidson (2002)	Futures options on S&P500, 9 interest rates, 5 currency, 4 energy, 3 metals, 10 agriculture 3 livestock	Various dates between Jan83-May2001	D	Implied $_{BK, NTM}$ $_{2Calls+2Puts\ equal\ weight}$ HIS_{30} GARCH	Overlapping option maturity, shortest but >10 days. Use sample SD of daily returns over forecast horizon to proxy "actual vol".	R^2 smaller for financial (23-28%), higher for metal & agriculture (30-37%), highest for livestock & energy (47,58%)	HIS30 and GARCH have little or no incremental information content. $\alpha_{implied}$ >0 for 24 cases (or 69%), all 35 cases $\beta_{implied}$ <1 with robust SE.
Taylor SJ (1986)	15 US stocks FT30 6 metal £/\$ 5 agricultural Futures 4 interest rate Futures	Jan66 – Dec 76 Jul75 – Aug82 Various length Nov74 – Sep 82 Various length Various length	D	EWMA Log-AR (1) ARMACH-Abs ARMACH-Sq HIS ARMACH-Sq is similar to GARCH	1 and 10 days ahead absolute returns. 2/3 of sample used in estimation. Use daily absolute returns deviation as "actual vol".	Relative MSE	Represents one of the earliest studies in ARCH class forecasts. The issue of volatility stationarity is not important when forecast over short horizon. The nonstationary series (e.g. EWMA) has the advantage of having fewer parameter estimates and forecasts respond to variance change fairly quickly.
Taylor SJ (1987)	DM/\$ futures	1977-1983	D	High, low and closing prices	1, 5, 10 & 20 days ahead. Estimation period, 5 years.	RMSE	The best model is a weighted average of present and past high, low and closing prices with adjustments for weekend and holiday effects.
Taylor SJ & Xu (1997)	DM/\$ DM/\$ options on PHLX	1/10/92 – 30/9/93 In: 9 months Out: 3 months	Quote D	Implied + ARCH combined Implied ARCH $HIS_{9\ months}$ $HIS_{last\ hour\ realised\ vol}$	1 hour ahead estimated from 9-month in-sample period. Use 5-min returns to proxy "actual vol".	MAE and MSE on std deviation & variance Friday macro news seasonal factors have no impact on forecast accuracy.	5-min return has information incremental to daily implied when forecasting hourly volatility. ARCH model includes with hourly and 5-min returns in the last hr plus 120 hour/day/week seasonal factors. Implied derived from NTM shortest maturity (>9 calendar days) Call+Put using BAW.

Thavaneswaran & Appadoo & Peiris (2005) (CA & AUS)		Theoretical Paper (N/A)		Stochastic volatility models: GARCH; Power GARCH; Non-stationary GARCH	L-steps-ahead	MMSE	The author plan to report their results in a future paper. (Thavaneswaran et al. / Statistics & Probability Letters)
Tims & Mahieu (2003)	USD/GBP USD/EUR USD/JPY GBP/JPY GBP/EUR JPY/EUR	1/9/89-22/7/02	D	The multivariate stochastic volatility model	Logarithmic high-low ranges of daily exchange rates as a proxy for volatility	Estimated measurement errors, transition errors, currency-specific components and LR test	Model fits the exchange rate data quite well.
Vilasuso (2002)	CAD/USD FRF/USD DEM/USD ITL/USD JPY/USD GBP/USD (Spot rate)	In: 13/3/79 – 31/12/97 Out: 1/1/98 – 31/12/99	D	FIGARCH GARCH, IGARCH (Ranked, GARCH marginally better than IGARCH)	1, 5 and 10 days ahead. Used daily squared returns to proxy actual volatility.	MSE, MAE, SE & Diebold-Mariano's test for sig. difference.	The FIGARCH model is better equipped to capture the salient features of exchange rate volatility than the commonly used GARCH and IGARCH models. And perhaps more important, the FIGARCH model generates superior out-of-sample volatility forecasts, and the gains in forecast accuracy are substantial. Significantly better forecasting performance from FIGARCH. Built FIARMA (with a constant term) on conditional variance without taking log. Truncated at lag 250.
Wei and Frankel (1991)	SrFr/\$ DM/\$ ¥/\$ £/\$ Options (PHLX)	2/83–1/90	M	Implied$_{\text{GK ATM call}}$ (Shortest maturity)	Non-overlapping 1 month ahead. Use sample SD of daily exchange rate return to proxy "actual vol"	R^2 30%(£), 17% (DM), 3%(SrFr), 0%(¥). $\alpha>0$, $\beta<1$ (except that for £/\$, $\alpha>0$, $\beta=1$) with heteroske consistent SE.	Use European formula for American style option. Also suffers from nonsynchronicity problem. Other tests reveal that Implied tends to over-predict high vol and under-predict low vol.
	Spot rates		D				Forecast/implied could be made more accurate by placing more weight on long run average.

West and Cho (1995)	C$/$ FFr/$ DM/$ ¥/$ £/$	14/3/73-20/9/89 In: 14/3/73 – 17/6/81 Out: 24/6/81 – 12/4/89	W	GARCH (1,1) IGARCH (1,1) AR (12) in absolute AR (12) in squares Homoskedastic Gaussian kernel	J=1, 12, 24 weeks estimated from rolling 432 weeks. Use J period squared returns to proxy actual volatility.	RMSPE and regression test on variance, R^2 varies from 0.1% to 4.5%.	GARCH models tend to make slightly more accurate forecasts. For longer horizons, it is difficult to find grounds for choosing between the various models. None of the models perform well in a conventional test of forecast efficiency. For a one-week horizon, GARCH models tend to make slightly more accurate forecasts. For longer horizons (12-24 weeks), it is difficult to find grounds for choosing between the various models. None of the models perform well in a conventional test of forecast efficiency. The nonparametric method came out worst though statistical tests for do not reject null of no significance difference in most cases.
Wright & Bollerslev (1999)	DEM/USD Spot rate	1/12/1986-1/12/1996	D Intraday (5-min)	GARCH (p, q) EGARCH (1, 1) AR (10)	One day ahead from rolling 288 5-minute observations in a day	MSE, R^2, coefficient estimates, AIC, SIC, or SE	Simple parametric GARCH models produce better forecasts than just fitting an autoregression to the squared returns, when working with daily data. When working with high frequency intra-daily data, simply fitting an autoregression to log-squared, squared, or absolute returns works better for forecasting future volatility than standard GARCH and EGARCH models fitted either to daily or intraday data.
Xu & Taylor (1995)	GBP/USD DEM/USD JPY/USD SFR/USD (PHLX options & Spot FX) Corresponding Futures Rates	2/1/85-8/1/92 In: 1/85-11/89 Out: 18/10/89-4/2/92	D	$\text{Implied}_{\text{BAW NTM TS or Short}}$ $\text{GARCH}_{\text{Normal or GED}}$ $\text{HIS}_{\text{4 weeks}}$	Non-overlapping 4 weeks ahead, estimated from a rolling sample of 250 weeks daily data. Use cumulative daily squared returns to proxy "actual vol".	ME, MAE, RMSE. When α_{implied} is set equal to 0, β_{implied} =1 cannot be rejected.	The implied predictors are superior to historical predictors. Implied works best and is unbiased. Other forecasts have no incremental information. GARCH forecast performance not sensitive to distributional assumption about returns. The choice of implied predictor (term structure, TS, or short maturity) does not affect results.

Yang (2005)	DEM/USD DEM/GBP	2/1/80- 30/10/92	D	The semiparametric GARCH model; GJR Model; GARCH (1, 1)	The rule-of-thumb bandwidth at 3212 observations	MSE, WMSE, Prediction Error	The semiparametric volatility model outperforms the GJR model as well as the more commonly used GARCH (1; 1) model in terms of goodness-of-fit, and forecasting, by correcting overgrowth in volatility.
Zumbach (2002)	USD/CHF USD/JPY	1/1/89 – 1/7/2000	H	LM-ARCH F-GARCH GARCH And their integrated counterparts	1 day ahead estimated from previous 5.5 years	RMSE. Realized volatility measured using hourly returns.	LM-ARCH, aggregates high frequency squared returns with a set of power law weights, is the best though difference is small. All integrated versions are more stable across time.
Zumbach, Pictet & Masutti	USD/CHF USD/JPY	1/1/1987- 31/12/1999 In: 1/1/1990- 31/12/1994 Out: 1/1/1995- 31/12/1999	H	Genetic Programming (GP); ARCH models: GARCH (1, 1), FIGARCH HARCH	One day ahead at daily interval (Time horizons ranging from 1 hour to a few weeks.)	RMSE	The out-of-sample forecasting performance of these new models are compared with the corresponding performance of some popular ARCH-types models and GP consistently outperforms the benchmarks. GP discovered that cross products of returns at different time horizons improve substantially the forecasting performance.

ABBREVIATION

A:
AIC and SIC: The Akaike and Schwartz information criteria
ANN: Artificial Neural Network
AP: Averaged Periodogram
AR-rv: Autoregressive Models of the Realized Variance (autoregressive projection on past realized volatilities)
ARCH: Autoregressive Conditional Heteroscedastic
ARFIMA: Autoregressive Fractionally Integrated Moving Average
ARIMA: Autoregressive Integrated Moving Average
A-RLS: Absolute Restricted Least Squares
ASE: asymptotic standard error
ATM: At the money
AUD: Australian Dollar
AVE: simple averaging (the model that produces a combined volatility forecast by averaging the forecasts of the four previous models)
B:
BAW: Barone-Adesi and Whaley American option pricing formula
BEF: Belgian Franc
BEKK Model: An earlier working paper by Baba, Engle, Kraft, and Kroner (Baba et. al. 1992)
BDS test: The BDS test developed by Brock et al. (1987) and generalized by Savit and Green (1991) and Wu et al. (1993) is aimed at distinguishing between data that are independently and identically distributed (iid) and data that reveal deterministic or stochastic dependence.
BDSL: BDSL test of Brock et al. (1996)
BS: Black-Scholes
C:
CAD: Canadian Dollar
CC-GARCH: constant correlation GARCH
CHF: Swiss Franc
CLCL: Close-to-Close volatility estimator
CVM: Cramer-von Mises statistic
D:
D: daily
DEM: German Mark
DKK: Danish Kroner
DW Test: Durbin Watson test
E:
EGARCH: Exponential GARCH
EMS: European Monetary System
ESP: Spanish Peseta
ESTAR: Exponential Smooth Transition Autoregressive (ESTAR) Model
EU: European Union
EUR: Euro
EWMA: Exponentially Weighted Moving Average
Exact ML: the estimator that maximizes the exact likelihood of log absolute returns, evaluated by simulation methods.

F:
FFF: Flexible Fourier Form
FIM: Finnish Markka
FIGARCH: Fractionally Integrated GARCH
FIV: Implied Volatilities
FRF: French Franc
FX Rates: Foreign Exchange Rates
G:
GARCH: Generalized ARCH
GBP: British Pound
GED: the Generalized Error Distribution
GEN:
GK: Garman and Kohlhagan model for pricing European currency option; Garman-Klass estimator
GRD: Greek Drachma
GW: Geometric Weight
H:
H: hourly
HARCH: Heterogeneous Autoregressive Conditional Heteroscedasticity
HAR-RV Model: The new forecasting model incorporating the jumps builds directly on the reduced form heterogenous AR model for the realized volatility.
HAR-RV-CJ Models: On defining the normalized multi-period jump and continuous sample path variability measures, the model obviously nests the HAR-RV-J model.
HAR-RV-J Model: HAR-RV forecasting model for the one-day volatilities extends straightforwardly to models for the realized volatilities over other horizons. Given the separate measurements of the jump components discussed above, these are readily included as additional explanatory variables over and above the longer-run realized volatility components, resulting in the new HAR-RV-J model.
His: Historical volatility constructed based on past variance/standard deviation.
HILO: High-Low estimator, the trading range estimator
HISVOL: Historical Volatility
HJM: Heath, Jarrow and Morton (1992) forward rate model for interest rates
HMAE: Heteroscedasticity Mean Absolute Error. HMAE is heteroskedasticity adjusted error statistics
HMSE: a Heteroscedasticity-adjusted version of MSE
HRMSE: Heteroscedasticity Root Mean Squared Errors. HRMSE is heteroskedasticity adjusted error statistics
I:
IEP: Irish Pound
IGARCH: Integrated GARCH
Implied: Implied Correlation
ISD: option implied standard deviation, based on the Black-Scholes model and various generalizations.
ITL: Italian Lira
IV: Implied Variance
J:
JPY: Japanese Yen

L:
LE: Logarithmic Error
LINEX: See p. 19, Poon & Granger (2002)
LL: Logarithmic Loss function.
LM: Lagrange Multiplier
LM-ARIMA: **long-memory ARIMA**
LP: Log Periodogram
LR: Likelihood Ratio test
LSD: Linear Squared Deviation
LUF: Luxembourg Franc
LW: Local Whittle
M:
M: monthly
MAD: mean absolute deviation
MAE: Mean Absolute Error
MA'E: Mean Average Error
MAFE: Mean Absolute Forecast Error
MAPE: Mean Absolute Percentage Error. MAPE measures percentage error against the "actual" and the comparison is made for each individual forecast error before aggregation.
MAV model: MAV is the MA model of volatility which forecasts volatility as the simple average of its own lagged values. (Pagan and Schwert (1990))
ME: Mean Error
MLAE: Mean Logarithm of Absolute Errors
MMAR: Multifractal Model of Asset Returns (MMAR)
MME: Mean Mixed Error
MME (U): under-predictions MME
MME (O): over-predictions MME
MMS: Money Market Services
MMSE: Minimum Mean Square Error
MSE: Mean Squared Error
MSM: Markov Switching Model
MS-GARCH: Markov Switching GARCH Models
N:
NLG: Dutch Guilder
NT: the new Taiwan dollar
NTM: Near the money
O:
OLS: Ordinary Least Squares
OTC: over-the-counter market
OTM: Out of the money
P:
PGARCH: Periodic GARCH
PHLX: Philadelphia Stock Exchange
PTE: Portuguese Escudo

Q:
Q: quarterly
OML: Quasi-Maximum Likelihood
QML (Log Abs. Return): the Gaussian quasi-maximum likelihood estimator with the log absolute return as volatility proxy
QML (Log Range): the Gaussian quasi-maximum likelihood estimator with the log range as volatility proxy
R:
R^2: Coefficients of determination
RID: Real Interest Differential Model
RiskMetrics-EWMA: RiskMetrics' Exponentially Weighted Moving Average Correlation
RLS: Restricted Least Squares
RMAE: Root Mean Absolute Error
RMSE: Root Mean Squared Errors
RMSFE: Root Mean Squared Forecast Error
RS: Regime Switching
RV: realized variance
S:
SFR: Swiss Franc
STD: standard deviation
SV: stochastic volatility model forecasts
SVM: Support Vector Machines
SVR: Support Vector Regression
T:
Theil-U Statistic: See p. 19, Poon & Granger (2002). Theil-U measure is similar to MAPE and compares model forecast error against a benchmark model forecast error and the ratio is derived at the aggregate level only.
U:
USD: US Dollar
W:
W: weekly
WACD: Weibull ACD
WLS: an implied volatility weighting scheme used in Whaley (1982) designed to minimize the pricing errors of a collection of options. In some cases the pricing errors are multiplied by trading volume or vega to give ATM implied a greater weight.
WMSE: Weighted Mean Square prediction Error
WRMSE: Weighted Root Mean Squared Errors

Appendix B.1

B.1 Calculating Innovations

In order to calculate innovations $\widehat{\zeta}_t$ in Equation (3.29), the key is to get a convenient expression of $(1+\theta_1 L+\theta_2 L^2)^{-1}$. The undetermined coefficient method in Method 1 is recommended for inverting the MA(2) in terms of numerical accuracy.

Method 1

The undetermined coefficient method is employed to calculate $\widehat{\zeta}_t$.

Factoring the polynomial $(1+\theta_1 L+\theta_2 L^2)$ as

$$(1+\theta_1 L+\theta_2 L^2) = (1-\lambda_1 L)(1-\lambda_2 L) = \left[1-(\lambda_1+\lambda_2)L+\lambda_1\lambda_2 L^2\right]$$

where

$$\lambda_1+\lambda_2 = -\theta_1$$

$$\lambda_1\lambda_2 = \theta_2$$

$$\because |\theta_1|>1\ ,\ \ |\theta_2|<1\ ,\ \ |\lambda_1|<1\ ,\ \ |\lambda_2|<1$$

$$\begin{aligned}
\therefore\ \ & (1+\theta_1 L+\theta_2 L^2)^{-1} \\
= \ \ & (1-\lambda_1 L)^{-1}(1-\lambda_2 L)^{-1} \\
= \ \ & \left[1+\lambda_1 L+\lambda_1^2 L^2+\lambda_1^3 L^3+...\right]\left[1+\lambda_2 L+\lambda_2^2 L^2+\lambda_2^3 L^3+...\right] \\
= \ \ & 1+(\lambda_1+\lambda_2)L+\left(\lambda_1^2+\lambda_1\lambda_2+\lambda_2^2\right)L^2+\left(\lambda_1^3+\lambda_1^2\lambda_2+\lambda_1\lambda_2^2+\lambda_2^3\right)L^3+...
\end{aligned}$$

Setting $K_1=\lambda_1+\lambda_2$, $K_2=\lambda_1^2+\lambda_1\lambda_2+\lambda_2^2$, $K_3=\lambda_1^3+\lambda_1^2\lambda_2+\lambda_1\lambda_2^2+\lambda_2^3 ...$ and substituting K into the equation above reveals that

$$(1+\theta_1 L+\theta_2 L^2)^{-1} = (1-\lambda_1 L)^{-1}(1-\lambda_2 L)^{-1} = 1+K_1 L+K_2 L^2+K_3 L^3+...$$

Multiplying both sides by $(1-\lambda_1 L)(1-\lambda_2 L)$ then we have

$$\begin{aligned} & (1+\theta_1 L+\theta_2 L^2)^{-1}(1-\lambda_1 L)(1-\lambda_2 L) \\ = \ & (1-\lambda_1 L)^{-1}(1-\lambda_2 L)^{-1}(1-\lambda_1 L)(1-\lambda_2 L) \\ = \ & (1-\lambda_1 L)(1-\lambda_2 L)\left(1+K_1 L+K_2 L^2+K_3 L^3+...\right) \end{aligned}$$

$$\because (1+\theta_1 L+\theta_2 L^2) = (1-\lambda_1 L)(1-\lambda_2 L)$$

$$\therefore (1+\theta_1 L+\theta_2 L^2)^{-1}(1+\theta_1 L+\theta_2 L^2) = 1 = (1+\theta_1 L+\theta_2 L^2)\left(1+K_1 L+K_2 L^2+K_3 L^3+...\right)$$

$$\Longrightarrow 1 = 1+[K_1+\theta_1]L+[K_2+K_1\theta_1+\theta_2]L^2+...$$

$$s.t. \quad K_1+\theta_1=0,\ K_2+K_1\theta_1+\theta_2=0...$$

$$\Longrightarrow K_1=-\theta_1,\ K_2=\theta_1^2-\theta_2...$$

So we get the innovations as follows

$$\begin{aligned} \widehat{\zeta}_t &= (1+\theta_1 L+\theta_2 L^2)^{-1}\Delta s_t^f \\ &= \left[1-\theta_1 L+\left(\theta_1^2-\theta_2\right)L^2+...\right]\Delta s_t^f \end{aligned}$$

Method 2

The determined coefficient method is employed to calculate $\widehat{\zeta}_t$.

Setting $\kappa_1=-\theta_1$, $\kappa_2=-\theta_2$, we rewrite Equation (3.29)

$$\widehat{\zeta}_t = \left[1-\kappa_1 L-\kappa_2 L^2\right]^{-1}\Delta s_t^f$$

Factoring the polynomial $(1-\kappa_1 L-\kappa_2 L^2)$ as

$$(1-\kappa_1 L-\kappa_2 L^2) = (1-\lambda_1 L)(1-\lambda_2 L) = \left[1-(\lambda_1+\lambda_2)L+\lambda_1\lambda_2 L^2\right]$$

where

$$\begin{aligned} \lambda_1+\lambda_2 &= \kappa_1 = -\theta_1 \\ \lambda_1\lambda_2 &= -\kappa_2 = \theta_2 \end{aligned}$$

Calculating λ_1 and λ_2 from $\lambda^2 - \kappa_1 L - \kappa_2 = (\lambda - \lambda_1)(\lambda - \lambda_2) = 0$

$$\lambda_1 = \frac{\kappa_1 + \sqrt{\kappa_1^2 + 4\kappa_2}}{2}$$
$$\lambda_2 = \frac{\kappa_1 - \sqrt{\kappa_1^2 + 4\kappa_2}}{2}$$

Substituting θ_1 and θ_2 into the equations above,

$$\lambda_1 = \frac{-\theta_1 + \sqrt{\theta_1^2 - 4\theta_2}}{2}$$
$$\lambda_2 = \frac{-\theta_1 - \sqrt{\theta_1^2 - 4\theta_2}}{2}$$

So we have the innovations as follows

$$\begin{aligned}
\widehat{\zeta}_t &= \left[1 - \kappa_1 L - \kappa_2 L^2\right]^{-1} \Delta s_t^f \\
&= (1 - \lambda_1 L)^{-1} (1 - \lambda_2 L)^{-1} \Delta s_t^f \\
&= (\lambda_1 - \lambda_2)^{-1} \left\{ \frac{\lambda_1}{1 - \lambda_1 L} - \frac{\lambda_2}{1 - \lambda_2 L} \right\} \Delta s_t^f \\
&\because \ |\lambda_1| \prec 1 \ , \ |\lambda_2| \prec 1, \\
&= \left\{ \frac{\lambda_1}{\lambda_1 - \lambda_2} \left[1 + \lambda_1 L + \lambda_1^2 L^2 + \lambda_1^3 L^3 + ...\right] - \frac{\lambda_2}{\lambda_1 - \lambda_2} \left[1 + \lambda_2 L + \lambda_2^2 L^2 + \lambda_2^3 L^3 + ...\right] \right\} \Delta s_t^f \\
&= (c_1 + c_2) \Delta s_t^f + (c_1 \lambda_1 + c_2 \lambda_2) \Delta s_{t-1}^f + \left(c_1 \lambda_1^2 + c_2 \lambda_2^2\right) \Delta s_{t-2}^f + \left(c_1 \lambda_1^3 + c_2 \lambda_2^3\right) \Delta s_{t-3}^f + ... \\
&= \sum_{j=0}^{\infty} \left(c_1 \lambda_1^j + c_2 \lambda_2^j\right) \Delta s_{t-j}^f
\end{aligned}$$

where

$$c_1 = \frac{\lambda_1}{\lambda_1 - \lambda_2}$$
$$c_2 = \frac{-\lambda_2}{\lambda_1 - \lambda_2}$$

We summarize the 3-step calculating process for innovations $\widehat{\zeta}_t$ that are subject to conditional volatility.

Step 1: Calculating Δs_t^f by

$$\Delta s_t^f = (1 - \rho_\pi L)(1 - \phi L) \Delta s_t = \Delta s_t - (\phi + \rho_\pi) \Delta s_{t-1} + \rho_\pi \phi \Delta s_{t-2}$$

where ϕ, ρ_π and the simulated s are available.

Step 2: Estimating θ_1 and θ_2 by

$$\Delta s_t^f = \zeta_t + \theta_1 \zeta_{t-1} + \theta_2 \zeta_{t-2}$$

where Δs_t^f is of an MA(2).

Step 3: Calculating $\widehat{\zeta_t}$ by the undetermined coefficient method

$$\widehat{\zeta_t} = (1 + \theta_1 L + \theta_2 L^2)^{-1} \Delta s_t^f = \left[1 - \theta_1 L + \left(\theta_1^2 - \theta_2\right) L^2 + ...\right] \Delta s_t^f$$

A series of innovations $\widehat{\zeta_t}$ which we can apply ARCH class models' estimations to directly is available.

Appendix B.2

B.2 ARCH Estimates and Tests

We report ARCH estimates and tests as well as steps in model selection process for the simulated data in detail.

As known, for GARCH class models, the simplest conditional mean equation is

$$y_t = c + bx_t + \tau_t \qquad \tau_t \sim N(0, h_t^2)$$

where the dependent variable y_t is written as a function of a constant c and an exogenous variable x_t with an error term τ_t. b is the parameter of the exogenous variable. The error term τ_t follows a normal distribution with a mean of zero and a variance h_t^2. We run the regressions of returns on the constant (and other variables that are depending upon the properties of the dependant variables) to make innovations / residuals available that are subject to ARCH.

Four different conditional mean equations are employed to maximally capture the ARCH property that is implied by the theoretical model in Section 3.5.

$$\Delta s_t = k_0 + \varrho_t \qquad \varrho_t \backsim N(0, \sigma^2_{\varrho_t}) \tag{B.1}$$

where the dependant variable is the spot return Δs_t that is regressed on a constant. Equation (B.1) tests if the innovation of spot return Δs_t has time properties of conditional volatility.

$$\Delta s_t = a_0 + a_1 \Delta s_{t-1} + a_2 \Delta s_{t-2} + a_3 \varepsilon_{t-1} + a_4 \varepsilon_{t-2} + \varepsilon_t \qquad \varepsilon_t \backsim N(0, \sigma^2_{\varepsilon_t}) \tag{B.2}$$

where the dependant variable is the spot return Δs_t that is regressed on a constant and an ARMA(2,2) process due to its properties. Equation (B.2) tests if the innovation of Δs_t that is an ARMA(2,2) implied by the model has time properties of conditional volatility.

$$\Delta s^f_t = b_0 + b_1 \xi_{t-1} + b_2 \xi_{t-2} + \xi_t \qquad \xi_t \backsim N(0, \sigma^2_{\xi_t}) \tag{B.3}$$

where the dependant variable is the filtered spot return Δs^f_t that is regressed on a constant and an MA(2) process due to the model's implied properties. Equation (B.3) tests if the innovation of Δs^f_t that is an MA(2) obtained by removing AR components of the ARMA(2,2) Δs_t has time properties of conditional volatility.

$$\widehat{\zeta}_t = g_0 + \mu_t \qquad \mu_t \backsim N(0, \sigma^2_{\mu_t}) \tag{B.4}$$

where the dependant variable is the innovation $\widehat{\zeta}_t$ that is regressed on a constant. Equation (B.4) tests if the innovation $\widehat{\zeta}_t$ that is produced by Equation (3.29) with the estimated θ_1 and θ_2 has time properties of conditional volatility. As showed in Equation (B.1), (B.2), (B.3), and (B.4), k_0, a_0, b_0 and g_0 are constants; ϱ_t, ε_t, ξ_t and μ_t are innovations. Innovations in the process are subject to ARCH. Now we refer τ_t as innovation (ϱ_t, ε_t, ξ_t and μ_t) and $h_t^2 (= \sigma^2_{\tau_t})$ as its conditional variance ($\sigma^2_{\varrho_t}$, $\sigma^2_{\varepsilon_t}$, $\sigma^2_{\xi_t}$, or $\sigma^2_{\mu_t}$).

To investigate conditional volatility, the best fitting ARCH class conditional volatility model to innovations needs to be estimated. The innovation τ_t produced in these four conditional mean equations mentioned above will be estimated and tested for "ARCH" properties for the standard GARCH class models (e.g. ARCH, GARCH, GARCH-M, EGARCH, TARCH, PARCH, and CGARCH). h_t^2 is the one-period ahead forecast variance based on past information and depends on the lag orders of squared errors, its previous own lags and other relevant variables determined by variants on the standard GARCH class models. h_t^2 not y_t is associated with ARCH. We report the results of ARCH effects and estimates for the theoretical residuals in baseline and sensitivity analysis respectively in Tables B1 - B60.

B.2.1 Residuals ϱ_t

Residuals ϱ_t are produced by Equation (B.1).

Before estimating the ARCH-type models, we use the Engle (1982) test for ARCH effects to make sure that this class of models is appropriate for the calculated residuals obtained by using OLS. Table B1 reports the test for ARCH effects. A test for the presence of ARCH in the residuals ϱ_t is calculated by regressing the squared residuals on a constant and p lags ($1 \leqslant p \leqslant 9$). Both the F-statistic and LM-statistic reported are highly significant (at 1%) in the baseline and sensitivity analysis for all lags (up to 9 lags). The innovations ϱ_t show the presence of ARCH effects.

We estimate GARCH type models on the squared residuals and test additional ARCH up to the order 9.

Symmetric conditional volatility models are reported in Tables B2 - B7.

Table B2 and Table B3 report estimates of ARCH(1) and tests for its additional ARCH. All coefficients in the conditional variance equation are highly statistically significant (at 1%). ARCH effects exist for all lags (1-9) in $\phi = 0.995$. There are some ARCH effects at the lags 5-9 in $\delta = -0.0025$ and at the lag 9 in the baseline.

We estimate the GARCH(1,1) model in Table B4 and test its additional ARCH up to order 9 in Table B5. In Table B4, the coefficients on all three terms in the conditional variance equation are highly significant (at 1%) in both the baseline and sensitivity analysis. There is no difference changing each parameter separately for sensitivity analysis in terms of significance. After the estimation of GARCH(1,1), there is the presence of ARCH left in its residuals. Specifically, the F-statistic and LM-statistic are very significant (at 1%) up to 9 lags in $\gamma = 0.7$, $\delta = -0.0025$ and $\phi = 0.995$ while ARCH effects also exist at the lag 1 (at 10%) in the baseline and at the lag 1 (at 5%) and 2 (at 10%) in $\rho_\pi = 0$.

GARCH-M suggests that the spot return Δs_t is determined by its risk because of the conditional variance of Δs_t in the conditional mean equation. The results of GARCH-M estimates and its ARCH tests found in Tables B6 and B7 are same to those of GARCH (1,1) reported in Tables B4 and B5. We find, but not report, that the parameters on the conditional variance term of returns in the mean equation are not significant in either the baseline or sensitivity analysis, which shows a non-significant relationship between return and risk.

Asymmetric conditional volatility models are reported in Tables B8 - B15.

In Table B8, we estimate the EGARCH(1,1) model. All of the ARCH and GARCH coefficients and constants are highly significant (at 1%) in $\gamma = 0.7$, $\delta = -0.0025$, $\rho_\pi = 0$ and $\phi = 0.995$ for sensitivity analysis. In the baseline, the ARCH term is significant (at 10%) while the GARCH term is not significant. The asymmetric term of EGARCH(1,1) is significant (at 10%) and negative (-0.087) only in $\rho_\pi = 0$, which says that the negative shocks imply a higher next period conditional variance than positive shocks of the same sign. We tests EGARCH(1,1) for additional ARCH up to order 9 in Table B9. ARCH effects are nearly highly significant for all lags except the non-significant lags (at the lag 1) in both the baseline and $\rho_\pi = 0$.

Threshold GARCH (TARCH) is reported in Tables B10 and B11. All of the constants, ARCH and GARCH coefficients in both the baseline and sensitivity analysis are highly significant (1%). Again, in $\rho_\pi = 0$, the asymmetric effects are captured. The coefficient estimate of the asymmetric term is significant (at 1%) and positive (+3.280), which suggests that news is asymmetric and bad news increases volatility. A leverage effect exists. In Table B11, TARCH(1,1) has exactly same ARCH effects compared to those found in GARCH(1,1) and GARCH-M.

Tables B12 and B13 display the PARCH(1,1)'s estimates and its additional ARCH tests. All coefficients on the ARCH and GARCH terms and the power parameters of the standard deviation in both the baseline and sensitivity analysis are highly significant (at 1%). All constant are significant. In $\phi = 0.995$ the asymmetric effects are captured with a significant (at 10%) and negative (-0.081) parameter. It is also found that volatility shocks are quite persistent in $\gamma = 0.7$ and $\delta = -0.0025$ for the sensitivity analysis because the sum of coefficients on the lagged squared error (ARCH term) and the lagged conditional variance (GARCH term) is very close to one where we define that the value of the sum of the ARCH and GARCH terms is on the interval [0.95, 1.05]. The presence of additional ARCH effects is in $\phi = 0.995$ up to the first four lags. In $\gamma = 0.7$ and $\rho_\pi = 0$, the highly significant ARCH effects (at 1%) are found up to lags 9 while none in the baseline. In $\delta = -0.0025$, additional ARCH is present at all lags (up to 9 lags) where additional ARCH is highly significant (at 1%) at the last seven lags (3-9) while it is significant (at 5%) at the first two lags.

The estimates and ARCH effects for the symmetric CGARCH(1,1) model are reported in Tables B14a and 15a. For the components in the permanent equation, the coefficients on the powers and the difference term of ARCH term minus GARCH term are highly significant (at 1%) with non-significant constants in both the baseline and sensitivity analysis; for the components in the transitory equation, all GARCH coefficients are highly significant (at 1%) while the ARCH terms are significant only in $\gamma = 0.7$ (at 10%) and in $\rho_\pi = 0$ (at 1%). $\rho_\pi = 0$ captures highly persistent shocks to the conditional variance. There is no additional ARCH effects found after estimating the symmetric CGARCH model.

Tables B14b and 15b report the asymmetric CGARCH(1,1) model's estimates and tests. For both the baseline and sensitivity analysis, all constants are not significant, and all parameters on the powers and the difference terms are highly significant (at 1%). The asymmetric CGARCH(1,1) term captures asymmetric effects with the significant (at 10% and at 1%) and positive (+0.059 and +0.068) coefficients in $\gamma = 0.7$ and $\phi = 0.995$, respectively. In the transitory equation, both the ARCH and GARCH estimates are significant in $\gamma = 0.7$ and $\phi = 0.995$ while neither in $\rho_\pi = 0$ and only the GARCH term is significant (at 5%) in both the baseline and $\delta = -0.0025$. Additional ARCH effects are not found after the asymmetric CGARCH estimates, which is same as those reported in Table B15a.

Testing and Estimating ARCH for the residuals ϱ_t which are produced by Equation (B.1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
Lag 1	^, "	^, "	^, "	^, "	^, "
Lag 2	^, "	^, "	^, "	^, "	^, "
Lag 3	^, "	^, "	^, "	^, "	^, "
Lag 4	^, "	^, "	^, "	^, "	^, "
Lag 5	^, "	^, "	^, "	^, "	^, "
Lag 6	^, "	^, "	^, "	^, "	^, "
Lag 7	^, "	^, "	^, "	^, "	^, "
Lag 8	^, "	^, "	^, "	^, "	^, "
Lag 9	^, "	^, "	^, "	^, "	^, "

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B1: ARCH effects in ϱ_t

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
C(2)	*	*	*	*	*
C(3)	*	*	*	*	*

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(2) - p-values of the constant coefficient in the variance equation of ARCH (1);
C(3) - p-values of the ARCH coefficient in the variance equation of ARCH (1).

Table B2: Estimates of ARCH(1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
Lag 1					^^^, """
Lag 2					^^, ""
Lag 3					^, "
Lag 4					^, "
Lag 5			^, "		^, "
Lag 6			^^, ""		^, "
Lag 7			^, "		^, "
Lag 8			^, "		^, "
Lag 9	^^, ""		^, "		^, "

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B3: ARCH effects in ARCH(1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
C(2)	*	*	*	*	*
C(3)	*	*	*	*	*
C(4)	*	*	*	*	*
C(3)+C(4)					

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(2) - p-values of the constant coefficient in the variance equation of GARCH (1,1);
C(3) - p-values of the ARCH coefficient in the variance equation of GARCH (1,1);
C(4) - p-values of the GARCH coefficient in the variance equation of GARCH (1,1).

Table B4: Estimates of GARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
Lag 1	^^^, """	^, "	^, "	^^, ""	^, "
Lag 2		^, "	^, "	^^^, """	^, "
Lag 3		^, "	^, "		^, "
Lag 4		^, "	^, "		^, "
Lag 5		^, "	^, "		^, "
Lag 6		^, "	^, "		^, "
Lag 7		^, "	^, "		^, "
Lag 8		^, "	^, "		^, "
Lag 9		^, "	^, "		^, "

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B5: ARCH effects in GARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
C(3)	*	*	*	*	*
C(4)	*	*	*	*	*
C(5)	*	*	*	*	*
C(4)+C(5)					

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(3) - p-values of the constant coefficient in the variance equation of GARCH-M;
C(4) - p-values of the ARCH coefficient in the variance equation of GARCH-M;
C(5) - p-values of the GARCH coefficient in the variance equation of GARCH-M.

Table B6: Estimates of GARCH-M

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
Lag 1	^^^ , """	^ , "	^ , "	^^ , ""	^ , "
Lag 2		^ , "	^ , "	^^^ , """	^ , "
Lag 3		^ , "	^ , "		^ , "
Lag 4		^ , "	^ , "		^ , "
Lag 5		^ , "	^ , "		^ , "
Lag 6		^ , "	^ , "		^ , "
Lag 7		^ , "	^ , "		^ , "
Lag 8		^ , "	^ , "		^ , "
Lag 9		^ , "	^ , "		^ , "

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B7: ARCH effects in GARCH-M

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
C(2)	*	*	*	*	*
C(3)	***	*	*	*	*
C(4)				(-0.087) ***	
C(5)		*	*	*	*
C(3)+C(5)					

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(2) - p-values of the constant coefficient in the variance equation of EGARCH (1,1);
C(3) - p-values of the ARCH coefficient in the variance equation of EGARCH (1,1);
C(4) - p-values of the asymmetry term in the variance equation of EGARCH (1,1);
C(5) - p-values of the GARCH coefficient in the variance equation of EGARCH (1,1).
(-) - the negative coefficient value.

Table B8: Estimates of EGARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
Lag 1		^ , "	^^ , ""		^ , "
Lag 2	^ , "	^ , "	^ , "	^ , "	^ , "
Lag 3	^ , "	^ , "	^ , "	^ , "	^ , "
Lag 4	^ , "	^ , "	^ , "	^ , "	^ , "
Lag 5	^ , "	^ , "	^ , "	^ , "	^ , "
Lag 6	^ , "	^ , "	^ , "	^ , "	^ , "
Lag 7	^ , "	^ , "	^ , "	^ , "	^ , "
Lag 8	^ , "	^ , "	^ , "	^ , "	^ , "
Lag 9	^ , "	^ , "	^ , "	^ , "	^ , "

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B9: ARCH effects in EGARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
C(2)	*	*	*	*	*
C(3)	*	*	*	*	*
C(4)				(+3.280) *	
C(5)	*	*	*	*	*
C(3)+C(5)					

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(2) - p-values of the constant coefficient in the variance equation of TARCH(1,1);
C(3) - p-values of the ARCH coefficient in the variance equation of TARCH(1,1);
C(4) - p-values of the threshold term in the variance equation of TARCH(1,1);
C(5) - p-values of the GARCH coefficient in the variance equation of TARCH(1,1).
(+) - the positive coefficient value.

Table B10: Estimates of TARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
Lag 1	^^^ , """	^ , "	^ , "	^^ , ""	^ , "
Lag 2		^ , "	^ , "	^^^ , """	^ , "
Lag 3		^ , "	^ , "		^ , "
Lag 4		^ , "	^ , "		^ , "
Lag 5		^ , "	^ , "		^ , "
Lag 6		^ , "	^ , "		^ , "
Lag 7		^ , "	^ , "		^ , "
Lag 8		^ , "	^ , "		^ , "
Lag 9		^ , "	^ , "		^ , "

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics

Table B11: ARCH effects in TARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
C(2)	**	*	*	***	*
C(3)	*	*	*	*	*
C(4)					(-0.081) ***
C(5)	*	*	*	*	*
C(6)	*	*	*	*	*
C(3)+C(5)		@	@		

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(2) - p-values of the constant coefficient in the std dev equation of PARCH(1,1);
C(3) - p-values of the ARCH coefficient in the std dev equation of PARCH(1,1);
C(4) - p-values of the asymmetry effects in the std dev equation of PARCH(1,1);
C(5) - p-values of the GARCH coefficient in the std dev equation of PARCH(1,1).
C(6) - p-values of the power parameter of the standard deviation in PARCH(1,1).
(-) - the negative coefficient value;
"@" - the sum of ARCH term and GARCH term is very close to unity, [0.95,1.05].

Table B12: Estimates of PARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
Lag 1		^ , "	^^ , ""	^ , "	^ , "
Lag 2		^ , "	^^ , ""	^ , "	^^ , ""
Lag 3		^ , "	^ , "	^ , "	^^^ , """
Lag 4		^ , "	^ , "	^ , "	^^^ , """
Lag 5		^ , "	^ , "	^ , "	
Lag 6		^ , "	^ , "	^ , "	
Lag 7		^ , "	^ , "	^ , "	
Lag 8		^ , "	^ , "	^ , "	
Lag 9		^ , "	^ , "	^ , "	

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B13: ARCH effects in PARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
C(2)					
C(3)	*	*	*	*	*
C(4)	*	*	*	*	*
C(5)		***		*	
C(6)	*	*	*	*	*
C(5)+C(6)				@	

At *(1%),**(5%),***(10%) significant levels for z-statistics;
C(2) - p-values of the constant coefficient in the long run component of CGARCH(1,1);
C(3) - p-values of powers in the long run component of CGARCH(1,1);
C(4) - p-values of the coefficient of the difference of (ARCH term - GARCH term) in the long run component of CGARCH(1,1);
C(5) - p-values of the ARCH coefficient in the transitory component of CGARCH(1,1);
C(6) - p-values of the GARCH coefficient in the transitory component of CGARCH(1,1);
"@" - the sum of ARCH term and GARCH term is very close to unity, [0.95,1.05].

Table B14a: Estimates of Sym. CGARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
C(2)					
C(3)	*	*	*	*	*
C(4)	*	*	*	*	*
C(5)		**			**
C(6)		(+0.059) ***			(+0.068) *
C(7)	**	*	**		*
C(5)+C(7)					

At *(1%),**(5%),***(10%) significant levels for z-statistics;
C(2) - p-values of the constant coefficient in the long run component of CGARCH(1,1);
C(3) - p-values of powers in the long run component of CGARCH(1,1);
C(4) - p-values of the coefficient of the difference of (ARCH term - GARCH term) in the long run component of CGARCH(1,1);
C(5) - p-values of the ARCH coefficient in the transitory component of CGARCH(1,1);
C(6) - p-values of the asymmetry coefficient in the transitory component of CGARCH(1,1);
C(7) - p-values of the GARCH coefficient in the transitory component of CGARCH(1,1);
(+) - the positive coefficient value.

Table B14b: Estimates of Asym.CGARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
Lag 1					
Lag 2					
Lag 3					
Lag 4					
Lag 5					
Lag 6					
Lag 7					
Lag 8					
Lag 9					

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B15a: ARCH effects in Sym. CGARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
Lag 1					
Lag 2					
Lag 3					
Lag 4					
Lag 5					
Lag 6					
Lag 7					
Lag 8					
Lag 9					

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B15b: ARCH effects in Asym. CGARCH(1,1)

B.2.2 Residuals ε_t

Residuals ε_t are generated by Equation (B.2).

Table B16 reports ARCH tests in the residuals ε_t that are calculated by regressing the squared residuals on a constant and up to 9 lags. Both the F-statistic and LM-statistic are highly significant (at 1%) in the baseline and sensitivity analysis for all lags (1 to 9). The innovations ε_t show the presence of ARCH effects.

We estimate GARCH type models and test additional ARCH up to the order 9.

Symmetric conditional volatility models are reported in Tables B17 - B22.

Tables B17 and B18 report ARCH(1)'s estimates and tests for its additional ARCH effects. All coefficients in the conditional variance equation are statistically significant where all constants and ARCH terms in sensitivity analysis are highly significant (at 1%) while the constant is highly significant (at 1%) and the ARCH term is significant (at 10%) in the baseline. ARCH effects exist at all lags (1-9) in $\phi = 0.995$, which is same to those in Table B3. ARCH effects are significant at most lags except at the lag 2 in $\delta = -0.0025$ and highly significant at the lags 4-9 in $\gamma = 0.7$ while no additional ARCH is found in the baseline and $\rho_\pi = 0$. In common for both Tables B3 and B18, there are no additional ARCH effects in $\rho_\pi = 0$.

We estimate the GARCH(1,1) model in Table B19 and test its additional ARCH up to order 9 in Table B20. In Table B19, the coefficients on all three terms in the conditional variance equation are highly significant (at 1%) in both the baseline and sensitivity analysis. There is no difference changing each parameter separately. After the estimates of GARCH(1,1), the presence of additional ARCH is found in the residuals. In details, the F-statistic and LM-statistic are very significant (at 1%) up to 9 lags in $\gamma = 0.7$, $\delta = -0.0025$ and $\phi = 0.995$ while ARCH effects also exist at the lag 1 (at 5%) and at the lag 2 (at 10%) in $\rho_\pi = 0$. Comparing Table B5 and Table B20, the only difference in the baseline is that there are no ARCH effects left in Table B20 while there are some that are significant (at 10%) at the lag 1 in Table B5.

GARCH-M investigates the relationship of the ARMA(2,2) spot returns Δs_t with the risk due to putting the conditional variance of Δs_t into the conditional mean equation. The results of GARCH-M' estimates and ARCH tests reported in Tables B21 and B22 are same as those of GARCH (1,1) reported in Tables B19 and B20. We find, but not report, that in the mean equation the parameters of the conditional variance term of returns are not significant in both the baseline and sensitivity analysis. The difference between Tables B7 and B22 is what it is between Tables B5 and B20 given above.

Asymmetric conditional volatility models are reported in Tables B23 - B30.

In Table B23, we estimate the EGARCH(1,1) model. All of the constants, ARCH and GARCH coefficients are highly significant (at 1%) for both the baseline and sensitivity analysis. In the baseline and $\rho_\pi = 0$ the asymmetric effects are captured. Both the asymmetric terms in the baseline and $\rho_\pi = 0$ are highly significant (at 1%) and negative (-0.049 and -0.055, respectively), which suggests that the negative shocks imply a higher next period conditional variance than positive shocks of the same sign. For the EGARCH(1,1) model, $\rho_\pi = 0$ capturing asymmetric effects is also found in Table B8. The results testing additional ARCH effects up to order 9 after estimating EGARCH(1,1) are displayed in Table B24. All of the F-statistic and LM-statistic are significant. The estimates and additional ARCH of the EGARCH model for the residuals ε_t in Tables B23 and B24 have better performance than those for the residuals ϱ_t in Tables B8 and B9. The EGARCH(1,1) model is a very good fitting model to conditional volatility for both the residuals.

The results of the TARCH model are reported in Tables B25 and B26. All of the constants, ARCH and GARCH coefficients are highly significant (at 1%), which is same with that reported in Table B10. Again, $\rho_\pi = 0$ captures asymmetric effects with a highly significant and positive parameter, which is found in Table B10 as well. In Table B25, the coefficient estimate of the asymmetric term in $\rho_\pi = 0$ is highly significant (at 1%) and positive (+2.947), which suggests that news is asymmetric and bad news increases volatility. There is a leverage effect. In Table B26, the same highly significant ARCH effects are found in $\gamma = 0.7$, $\delta = -0.0025$ and $\phi = 0.995$ as in Table B11. In the baseline, no ARCH is found in Table B26 while both the F-statistic and LM-statistic are significant (at 10%) at the lag 1 in Table B11. Additional ARCH in $\rho_\pi = 0$ is similar to that found in Table B11, where in Table B26 additional ARCH is significant at the first three lags (at 5% on the first two lags and at 10% on the lag 3) while in Table B11 additional ARCH is significant at the lag 1 (at 5%) and at the lag 2 (at 10%).

Tables B27 and B28 display the PARCH(1,1)'s estimates and additional ARCH tests. For both the baseline and sensitivity analysis, almost all of the ARCH, GARCH and the standard deviation power parameters are significant except those non-significant terms in $\delta = -0.0025$. Constant terms are significant in the baseline, $\gamma = 0.7$ and $\phi = 0.995$. Regarding the ability of capturing asymmetric effects, the asymmetric terms are significant in $\gamma = 0.7$ and $\rho_\pi = 0$. Both estimates of the parameter of the asymmetric terms are significant (at 5%): one is negative (-0.101) in $\gamma = 0.7$ and the other is positive (+0.293) in $\rho_\pi = 0$. It is also found that volatility shocks are quite persistent in $\gamma = 0.7$ and $\rho_\pi = 0$ for the sensitivity analysis due to the sum of the coefficients of the lagged squared error (ARCH term) and the lagged conditional variance (GARCH term) very close to one on the interval [0.95, 1.05]. In $\gamma = 0.7$ PARCH(1,1) is able to capture highly persistent volatility shocks and the presence of additional ARCH effects is highly significant (at 1%) up to lags 9, which is found as well previously in Table B12 and B13. Moreover, additional ARCH is also found at all lags in $\rho_\pi = 0$ and $\phi = 0.995$ while there is no additional ARCH effects found in the baseline and $\delta = -0.0025$. In Tables B27 and B28, $\gamma = 0.7$ dominates.

The estimates and ARCH tests for the symmetric CGARCH(1,1) model are reported in Tables B29a and B30a. In the permanent equation of the CGARCH model, the coefficients of the powers and the difference term are highly significant (at 1%) with non-significant constants for both the baseline and sensitivity analysis; in the transitory equation of the CGARCH model, both ARCH and GARCH terms are significant in the baseline and $\gamma = 0.7$ while both terms are not significant in $\delta = -0.0025$ and $\rho_\pi = 0$, and only the GARCH term is significant in $\phi = 0.995$. There are no additional ARCH effects after the CGARCH estimates, which is the same as that found in Table B15a.

Tables B29b and B30b report the results for the asymmetric CGARCH(1,1) model. In the permanent equation, all coefficients of the powers and the difference terms are highly significant (at 1 %) for both the baseline and sensitivity analysis while the constants are significant only in the baseline and $\rho_\pi = 0$. In the transitory equation, the GARCH coefficients are significant except the one in $\phi = 0.995$. The ARCH terms in the baseline and $\rho_\pi = 0$ are significant at 1% and at 5% significance level, respectively. The asymmetric CGARCH(1,1) model captures asymmetric effects in the baseline, $\delta = -0.0025$ and $\rho_\pi = 0$ where the asymmetric coefficients are significant (at 1%, 10%, 1%, respectively) with negative or positive signs (+0.234 in the baseline, -0.064 in $\delta = -0.0025$, and +0.107 in $\rho_\pi = 0$). Additional ARCH is found in the baseline where it is significant at the lags (2-7).

It follows that we find and report that, for the residuals generated by Equation (B.3) and Equation (B.4), the results of the ARCH tests and estimates are not only similar to but also suggest more information than those found in Equations (B.1) and (B.2).

Testing and Estimating ARCH for the residuals ε_t which are produced by Equation (B.2)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
Lag 1	^, "	^, "	^, "	^, "	^, "
Lag 2	^, "	^, "	^, "	^, "	^, "
Lag 3	^, "	^, "	^, "	^, "	^, "
Lag 4	^, "	^, "	^, "	^, "	^, "
Lag 5	^, "	^, "	^, "	^, "	^, "
Lag 6	^, "	^, "	^, "	^, "	^, "
Lag 7	^, "	^, "	^, "	^, "	^, "
Lag 8	^, "	^, "	^, "	^, "	^, "
Lag 9	^, "	^, "	^, "	^, "	^, "

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B16: ARCH effects in ε_t

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
C(6)	*	*	*	*	*
C(7)	***	*	*	*	*

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(6) - p-values of the constant coefficient in the variance equation of ARCH (1);
C(7) - p-values of the ARCH coefficient in the variance equation of ARCH (1).

Table B17: Estimates of ARCH(1)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
Lag 1			^^^, """		^^^, """
Lag 2					^^, ""
Lag 3			^, "		^, "
Lag 4		^, "	^, "		^, "
Lag 5		^, "	^, "		^, "
Lag 6		^, "	^, "		^, "
Lag 7		^, "	^, "		^, "
Lag 8		^, "	^, "		^, "
Lag 9		^, "	^, "		^, "

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B18: ARCH effects in ARCH(1)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
C(6)	*	*	*	*	*
C(7)	*	*	*	*	*
C(8)	*	*	*	*	*
C(7)+C(8)					

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(6) - p-values of the constant coefficient in the variance equation of GARCH (1,1);
C(7) - p-values of the ARCH coefficient in the variance equation of GARCH (1,1);
C(8) - p-values of the GARCH coefficient in the variance equation of GARCH (1,1).

Table B19: Estimates of GARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
Lag 1		^, "	^, "	^^, ""	^, "
Lag 2		^, "	^, "	^^^, """	^, "
Lag 3		^, "	^, "		^, "
Lag 4		^, "	^, "		^, "
Lag 5		^, "	^, "		^, "
Lag 6		^, "	^, "		^, "
Lag 7		^, "	^, "		^, "
Lag 8		^, "	^, "		^, "
Lag 9		^, "	^, "		^, "

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B20: ARCH effects in GARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
C(7)	*	*	*	*	*
C(8)	*	*	*	*	*
C(9)	*	*	*	*	*
C(8)+C(9)					

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(7) - p-values of the constant coefficient in the variance equation of GARCH-M;
C(8) - p-values of the ARCH coefficient in the variance equation of GARCH-M;
C(9) - p-values of the GARCH coefficient in the variance equation of GARCH-M.

Table B21: Estimates of GARCH-M

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
Lag 1		^, "	^, "	^^, ""	^, "
Lag 2		^, "	^, "	^^^, """	^, "
Lag 3		^, "	^, "		^, "
Lag 4		^, "	^, "		^, "
Lag 5		^, "	^, "		^, "
Lag 6		^, "	^, "		^, "
Lag 7		^, "	^, "		^, "
Lag 8		^, "	^, "		^, "
Lag 9		^, "	^, "		^, "

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B22: ARCH effects in GARCH-M

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
C(6)	*	*	*	*	*
C(7)	*	*	*	*	*
C(8)	(-0.049) *			(-0.055) *	
C(9)	*	*	*	*	*
C(7)+C(9)					

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(6) - p-values of the constant coefficient in the variance equation of EGARCH (1,1);
C(7) - p-values of the ARCH coefficient in the variance equation of EGARCH (1,1);
C(8) - p-values of the asymmetry term in the variance equation of EGARCH (1,1);
C(9) - p-values of the GARCH coefficient in the variance equation of EGARCH (1,1).
(-) - the negative coefficient value.

Table B23: Estimates of EGARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
Lag 1	^^^, """	^^, ""	^^, ""	^, "	^, "
Lag 2	^^, ""	^, "	^, "	^, "	^, "
Lag 3	^^^, """	^, "	^, "	^, "	^, "
Lag 4	^^^, """	^, "	^, "	^, "	^, "
Lag 5	^^, ""	^, "	^, "	^, "	^, "
Lag 6	^^^, """	^, "	^, "	^, "	^, "
Lag 7	^^^, """	^, "	^, "	^, "	^, "
Lag 8	^^^, """	^, "	^, "	^, "	^, "
Lag 9	^^^, """	^, "	^, "	^, "	^, "

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B24: ARCH effects in EGARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
C(6)	*	*	*	*	*
C(7)	*	*	*	*	*
C(8)				(+2.947) *	
C(9)	*	*	*	*	*
C(7)+C(9)					

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(6) - p-values of the constant coefficient in the variance equation of TARCH(1,1);
C(7) - p-values of the ARCH coefficient in the variance equation of TARCH(1,1);
C(8) - p-values of the threshold term in the variance equation of TARCH(1,1);
C(9) - p-values of the GARCH coefficient in the variance equation of TARCH(1,1).
(+) - the positive coefficient value.

Table B25: Estimates of TARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
Lag 1		^, "	^, "	^^, ""	^, "
Lag 2		^, "	^, "	^^, ""	^, "
Lag 3		^, "	^, "	^^^, """	^, "
Lag 4		^, "	^, "		^, "
Lag 5		^, "	^, "		^, "
Lag 6		^, "	^, "		^, "
Lag 7		^, "	^, "		^, "
Lag 8		^, "	^, "		^, "
Lag 9		^, "	^, "		^, "

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B26: ARCH effects in TARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
C(6)	*	*			*
C(7)	*	*		*	*
C(8)		(-0.101) **		(+0.293) **	
C(9)	*	*		*	*
C(10)	*	*		***	*
C(7)+C(9)		@		@	

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(6) - p-values of the constant coefficient in the std dev equation of PARCH(1,1);
C(7) - p-values of the ARCH coefficient in the std dev equation of PARCH(1,1);
C(8) - p-values of the asymmetry effects in the std dev equation of PARCH(1,1);
C(9) - p-values of the GARCH coefficient in the std dev equation of PARCH(1,1).
C(10) - p-values of the power parameter of the standard deviation in PARCH(1,1);
(+/-) - the positive / negative coefficient value.
@ - the sum of ARCH term and GARCH term is very close to unity, [0.95,1.05]

Table B27: Estimates of PARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
Lag 1		^, "		^, "	^, "
Lag 2		^, "		^, "	^, "
Lag 3		^, "		^, "	^, "
Lag 4		^, "		^, "	^, "
Lag 5		^, "		^, "	^^, ""
Lag 6		^, "		^, "	^^, ""
Lag 7		^, "		^, "	^^, ""
Lag 8		^, "		^, "	^^, ""
Lag 9		^, "		^, "	^^, ""

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B28: ARCH effects in PARCH(1,1)

CGARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
C(6)					
C(7)	*	*	*	*	*
C(8)	*	*	*	*	*
C(9)	*	**			
C(10)	*	*			*
C(9)+C(10)					

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(6) - p-values of the constant coefficient in the long run component of CGARCH(1,1);
C(7) - p-values of powers in the long run component of CGARCH(1,1);
C(8) - p-values of the coefficient of the difference of (ARCH term - GARCH term) in the long run component of CGARCH(1,1);
C(9) - p-values of the ARCH coefficient in the transitory component of CGARCH(1,1);
C(10) - p-values of the GARCH coefficient in the transitory component of CGARCH(1,1).

Table B29a: Estimates of Sym. CGARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
C(6)	*			*	
C(7)	*	*	*	*	*
C(8)	*	*	*	*	*
C(9)	*			**	
C(10)	(+0.234) *		(-0.064) ***	(+0.107) *	
C(11)	*	*	*	**	
C(9)+C(11)					

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(6) - p-values of the constant coefficient in the long run component of CGARCH(1,1);
C(7) - p-values of powers in the long run component of CGARCH(1,1);
C(8) - p-values of the coefficient of the difference of (ARCH term - GARCH term) in the long run component of CGARCH(1,1);
C(9) - p-values of the ARCH coefficient in the transitory component of CGARCH(1,1);
C(10) - p-values of the asymmetry coefficient in the transitory component of CGARCH(1,1
C(11) - p-values of the GARCH coefficient in the transitory component of CGARCH(1,1).
(+/-) - the positive / negative coefficient value.

Table B29b: Estimates of Asym.CGARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
Lag 1					
Lag 2					
Lag 3					
Lag 4					
Lag 5					
Lag 6					
Lag 7					
Lag 8					
Lag 9					

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B30a: ARCH effects in Sym. CGARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
Lag 1					
Lag 2	^^^, """				
Lag 3	^^^, """				
Lag 4	^^^, """				
Lag 5	^^, ""				
Lag 6	^^^, """				
Lag 7	^^^, """				
Lag 8					
Lag 9					

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B30b: ARCH effects in Asym. CGARCH(1,1)

We report the estimates of all coefficients in the conditional variance equation for the standard GARCH class models in terms of the significance of the coefficients, asymmetric effects, persistent shocks, and the relationship of return with risk.

B.2.3 Residuals ξ_t

Residuals ξ_t are produced by Equation (B.3).

Table B31 reports the presence of ARCH in the residuals ξ_t that is calculated by regressing the squared residuals on a constant and up to 9 lags. Both the F-statistic and LM-statistic are highly significant (at 1%) at all lags (1 to 9) in the baseline and sensitivity analysis. The innovations ξ_t show the evidence of ARCH effects.

We estimate the GARCH type models and test additional ARCH up to order 9.

Symmetric conditional volatility models are reported in Tables B32 - B37.

Tables B32 and B33 report ARCH estimates and tests for the ARCH(1) model. All coefficients in the conditional variance equation are significant in the baseline and sensitivity analysis. ARCH effects are significant for all lags (1-9) in $\phi = 0.995$, which is similar with those found in Tables B3 and B18. There are significant ARCH effects at the lags 3-9 in $\gamma = 0.7$ and at the lags 4-9 in $\delta = -0.0025$. In Table B33, no ARCH effect is found in the baseline, which is same to that as in Table B18, and in $\rho_\pi = 0$, which is same to those as in Table B3 and Table B18.

We estimate the GARCH(1,1) model in Table B34 and test additional ARCH up to order 9 in Table B35. In Table B34, the coefficients on all three terms in the conditional variance equation are highly significant (at 1%) in the baseline and sensitivity analysis. There is no difference changing each parameter separately. After the estimations of GARCH(1,1), ARCH effects are found in the residuals. Specifically, the F-statistic and LM-statistic are significant at the lag 1 in the baseline and at all lags (up to 9 lags) in the sensitivity analysis. Comparing Tables B5, B20 and B35, it is found that, in common, ARCH effects are present at all lags in $\gamma = 0.7$, $\delta = -0.0025$ and $\phi = 0.995$.

The GARCH-M model investigates if the filtered spot return Δs_t^f is determined by its risk in terms of the conditional variance of Δs_t^f in the conditional mean equation. The estimates and tests of GARCH-M are reported in Tables B36 and B37. All constants, ARCH and GARCH terms in the conditional variance equation are highly significant (at 1%), which is same as those in Tables B6 and B21. Additional ARCH effects of GARCH-M exist not only significantly at lags 1-2 in the baseline and the lags 7-9 in $\rho_\pi = 0$ but also highly significantly (at 1%) at all lags (1-9) in $\gamma = 0.7$, $\delta = -0.0025$ and $\phi = 0.995$. We find, but do not report, that in the mean equation the parameters of the conditional variance term of returns are positive and highly significant (at 1%) in the baseline, $\gamma = 0.7$ and $\delta = -0.0025$ and significant (at 5%) in $\rho_\pi = 0$. It suggests that the filtered spot return Δs_t^f is partly determined by its risk.

Asymmetric conditional volatility models are reported in Tables B38 - B45.

In Table B38, the EGARCH(1,1) model is estimated. All of the constants, ARCH and GARCH coefficients in the conditional variance equation are highly significant (at 1%), which is same as that reported in Table B23. The EGARCH(1,1) model can capture the asymmetric effects well for the baseline and sensitivity analysis except in $\rho_\pi = 0$. The estimates of the asymmetric parameters are significant (at 1% in both the baseline and $\phi = 0.995$; at 5% in both $\gamma = 0.7$ and $\delta = -0.0025$). Three out of four significant asymmetric parameters are negative (-0.074 in the baseline, -0.028 in $\gamma = 0.7$ and -0.033 in $\delta = -0.0025$) and the rest one is positive (+0.047 in $\phi = 0.995$). The additional ARCH of EGARCH(1,1) is displayed in Table B39. ARCH effects are highly significant (at 1%) at all lags (1-9) in $\rho_\pi = 0$, which is same to that in Table B24, and significant at lags 6-9 in $\gamma = 0.7$, at the lags 5-9 in $\delta = -0.0025$, and at the lag 2 in $\phi = 0.995$. The EGARCH(1,1) model for the residuals ξ_t in Table B38 has better performance capturing asymmetric effects than it does in Tables B8 and B23 for the residuals ϱ_t and ε_t, respectively.

The TARCH model is reported in Tables B40 and B41. All of the constants, ARCH and GARCH coefficients are highly significant (at 1%), which is same with those reported in Tables B10 and B25. In the baseline, $\delta = -0.0025$ and $\phi = 0.995$ the asymmetric coefficients are significant (at 10%, 10%, and 5%, respectively) and positive (+0.266, +0.102, and +0.124 respectively) while the significant and positive asymmetric coefficients exist in $\rho_\pi = 0$ in Tables B10 and B25, suggesting that news is asymmetric and bad news increases volatility with a leverage effect. There are almost same additional ARCH effects in Table B41 with those to GARCH in Table B35. Comparing Tables B11, B26 and B41 for the TARCH(1,1) model, it is common that additional ARCH effects are significant in $\gamma = 0.7$, $\delta = -0.0025$ and $\phi = 0.995$.

Tables B42 and B43 display the PARCH(1,1)'s estimates and ARCH tests. All of the ARCH and GARCH coefficients and the power parameters of the standard deviation are highly significant (at 1%) with significant constants in both the baseline and sensitivity analysis. The asymmetric coefficient is significant (at 10%) and positive (+0.093) only in $\phi = 0.995$ where volatility shocks are quite persistent. The presence of additional ARCH effects of PARCH(1,1) is significant (at 10%) at the lag 9 in $\gamma = 0.7$, at the lags 6-7 in $\rho_\pi = 0$, and at the lag 1 in $\phi = 0.995$. There are no ARCH effects found in the baseline and $\delta = -0.0025$, which is same with those reported in Table B28. In common, there is no additional ARCH in the baseline as in Tables B13, B28, and B43.

The ARCH estimates and tests for the symmetric CGARCH(1,1) are reported in Tables B44a and B45a. All coefficients except the non-significant constants in the permanent equation are significant; in the transitory equation, only in $\gamma = 0.7$ both the ARCH and GARCH terms are significant, and in $\phi = 0.995$ only the GARCH (ARCH) term is not significant (significant at 5%) while others on the GARCH term are significant (at 1%, 5%, and 1% in the baseline, $\delta = -0.0025$ and $\rho_\pi = 0$, respectively). In $\gamma = 0.7$ CGARCH(1,1) captures highly persistent shocks to the conditional variance. Additional ARCH effects of the symmetric CGARCH are present at the lag 2 in $\rho_\pi = 0$ and at the lags 1-7 in $\phi = 0.995$ while no ARCH effects are found in the baseline, $\gamma = 0.7$, and $\delta = -0.0025$.

Tables B44b and B45b report the estimates and tests for the asymmetric CGARCH(1,1) model. In the permanent equation, all coefficients of the powers and the difference terms are highly significant (at 1%) in the both baseline and sensitivity analysis. The constants are significant in the baseline, $\delta = -0.0025$ and $\rho_\pi = 0$. The ARCH coefficients are significant in the baseline (at 10%) and $\rho_\pi = 0$ (at 1%). All GARCH terms are significant in the baseline (at 1%) and sensitivity analysis (at 5%, 1%, 1%, and 1% in $\gamma = 0.7$, $\delta = -0.0025$, $\rho_\pi = 0$, and $\phi = 0.995$, respectively). Asymmetric effects are captured by the asymmetric CGARCH in the baseline, $\delta = -0.0025$ and $\rho_\pi = 0$, which is found same with those in Table B29b. The asymmetric parameters in the baseline, $\delta = -0.0025$ and $\rho_\pi = 0$ in Table B44b are significant (at 1%, 10%, and 1%, respectively) with a positive (+0.343), negative (-0.039) and positive (+0.253) signs, respectively. In Table B45b, additional ARCH effects are found not only at the lag 2 in the baseline and at the lag 9 in $\gamma = 0.7$ but also at all lags (1-9) in $\rho_\pi = 0$, which is compared to those in Tables B15b and B30b where no additional ARCH exists in $\gamma = 0.7$ and $\rho_\pi = 0$. Overall, the asymmetric CGARCH model does well for capturing asymmetric effects in Tables B14b, B29b, and B44b.

Testing and Estimating ARCH for the residuals ξ_t which are produced by Equation (B.3)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
Lag 1	^, "	^, "	^, "	^, "	^, "
Lag 2	^, "	^, "	^, "	^, "	^, "
Lag 3	^, "	^, "	^, "	^, "	^, "
Lag 4	^, "	^, "	^, "	^, "	^, "
Lag 5	^, "	^, "	^, "	^, "	^, "
Lag 6	^, "	^, "	^, "	^, "	^, "
Lag 7	^, "	^, "	^, "	^, "	^, "
Lag 8	^, "	^, "	^, "	^, "	^, "
Lag 9	^, "	^, "	^, "	^, "	^, "

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B31: ARCH effects in ξ_t

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
C(4)	*	*	*	*	*
C(5)	*	**	*	*	*

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(4) - p-values of the constant coefficient in the variance equation of ARCH (1);
C(5) - p-values of the ARCH coefficient in the variance equation of ARCH (1).

Table B32: Estimates of ARCH(1)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
Lag 1					^^, ""
Lag 2					^, "
Lag 3		^, "			^, "
Lag 4		^, "	^^^, """		^, "
Lag 5		^, "	^^^, """		^, "
Lag 6		^, "	^^^, """		^, "
Lag 7		^, "	^, "		^, "
Lag 8		^, "	^, "		^, "
Lag 9		^, "	^, "		^, "

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B33: ARCH effects in ARCH(1)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
C(4)	*	*	*	*	*
C(5)	*	*	*	*	*
C(6)	*	*	*	*	*
C(5)+C(6)					

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(4) - p-values of the constant coefficient in the variance equation of GARCH (1,1);
C(5) - p-values of the ARCH coefficient in the variance equation of GARCH (1,1);
C(6) - p-values of the GARCH coefficient in the variance equation of GARCH (1,1).

Table B34: Estimates of GARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
Lag 1	^^^, """	^, "	^, "	^, "	^^, ""
Lag 2		^, "	^, "	^^, ""	^, "
Lag 3		^, "	^, "	^^, ""	^, "
Lag 4		^, "	^, "	^^, ""	^, "
Lag 5		^, "	^, "	^^, ""	^, "
Lag 6		^, "	^, "	^, "	^, "
Lag 7		^, "	^, "	^, "	^, "
Lag 8		^, "	^, "	^, "	^, "
Lag 9		^, "	^, "	^^, ""	^, "

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B35: ARCH effects in GARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
C(5)	*	*	*	*	*
C(6)	*	*	*	*	*
C(7)	*	*	*	*	*
C(6)+C(7)					

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(5) - p-values of the constant coefficient in the variance equation of GARCH-M;
C(6) - p-values of the ARCH coefficient in the variance equation of GARCH-M;
C(7) - p-values of the GARCH coefficient in the variance equation of GARCH-M.

Table B36: Estimates of GARCH-M

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
Lag 1	^^^, """	^, "	^, "		^, "
Lag 2	^^^, """	^, "	^, "		^, "
Lag 3		^, "	^, "		^, "
Lag 4		^, "	^, "		^, "
Lag 5		^, "	^, "		^, "
Lag 6		^, "	^, "		^, "
Lag 7		^, "	^, "	^^, ""	^, "
Lag 8		^, "	^, "	^^, ""	^, "
Lag 9		^, "	^, "	^^^, """	^, "

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B37: ARCH effects in GARCH-M

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
C(4)	*	*	*	*	*
C(5)	*	*	*	*	*
C(6)	(-0.074) *	(-0.028) **	(-0.033) **		(+0.047) *
C(7)	*	*	*	*	*
C(5)+C(7)					

At *(1%),**(5%),***(10%) significant levels for z-statistics;
C(4) - p-values of the constant coefficient in the variance equation of EGARCH (1,1);
C(5) - p-values of the ARCH coefficient in the variance equation of EGARCH (1,1);
C(6) - p-values of the asymmetry term in the variance equation of EGARCH (1,1);
C(7) - p-values of the GARCH coefficient in the variance equation of EGARCH (1,1).
(+/-) - the positive / negative coefficient value.

Table B38: Estimates of EGARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
Lag 1				^, "	
Lag 2				^, "	^^^, """
Lag 3				^, "	
Lag 4				^, "	
Lag 5			^^, ""	^, "	
Lag 6		^^^, """	^, "	^, "	
Lag 7		^, "	^, "	^, "	
Lag 8		^, "	^, "	^, "	
Lag 9		^, "	^^, ""	^, "	

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B39: ARCH effects in EGARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
C(4)	*	*	*	*	*
C(5)	*	*	*	*	*
C(6)	(+0.266) ***		(+0.102) ***		(+0.124) **
C(7)	*	*	*	*	*
C(5)+C(7)					

At *(1%),**(5%),***(10%) significant levels for z-statistics;
C(4) - p-values of the constant coefficient in the variance equation of TARCH(1,1);
C(5) - p-values of the ARCH coefficient in the variance equation of TARCH(1,1);
C(6) - p-values of the threshold term in the variance equation of TARCH(1,1);
C(7) - p-values of the GARCH coefficient in the variance equation of TARCH(1,1).
(+) - the positive coefficient value.

Table B40: Estimates of TARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
Lag 1	^^^, """	^, "	^, "	^, "	^, "
Lag 2		^, "	^, "	^^, ""	^, "
Lag 3		^, "	^, "	^^, ""	^, "
Lag 4		^, "	^, "	^^, ""	^, "
Lag 5		^, "	^, "	^^, ""	^, "
Lag 6		^, "	^, "	^, "	^, "
Lag 7		^, "	^, "	^, "	^, "
Lag 8		^, "	^, "	^, "	^, "
Lag 9		^, "	^, "	^^, ""	^, "

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B41: ARCH effects in TARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
C(4)	**	*	*	**	*
C(5)	*	*	*	*	*
C(6)					(+0.093) ***
C(7)	*	*	*	*	*
C(8)	*	*	*	*	*
C(5)+C(7)					@

At *(1%),**(5%),***(10%) significant levels for z-statistics;
C(4) - p-values of the constant coefficient in the std dev equation of PARCH(1,1);
C(5) - p-values of the ARCH coefficient in the std dev equation of PARCH(1,1);
C(6) - p-values of the asymmetry effects in the std dev equation of PARCH(1,1);
C(7) - p-values of the GARCH coefficient in the std dev equation of PARCH(1,1).
C(8) - p-values of the power parameter of the standard deviation in PARCH(1,1);
(+) - the positive coefficient value;
@ - the sum of ARCH term and GARCH term is very close to unity, [0.95,1.05].

Table B42: Estimates of PARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
Lag 1					^^^ , """
Lag 2					
Lag 3					
Lag 4					
Lag 5					
Lag 6				^^^ , """	
Lag 7				^^^ , """	
Lag 8					
Lag 9		^^^ , """			

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B43: ARCH effects in PARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
C(4)					
C(5)	*	*	*	*	*
C(6)	*	*	*	**	*
C(7)		***			**
C(8)	*	*	**	*	
C(7)+C(8)		@			

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(4) - p-values of the constant coefficient in the long run component of CGARCH(1,1);
C(5) - p-values of powers in the long run component of CGARCH(1,1);
C(6) - p-values of the coefficient of the difference of (ARCH term - GARCH term) in the long run component of CGARCH(1,1);
C(7) - p-values of the ARCH coefficient in the transitory component of CGARCH(1,1);
C(8) - p-values of the GARCH coefficient in the transitory component of CGARCH(1,1).
@ - the sum of ARCH term and GARCH term is very close to unity, [0.95,1.05].

Table B44a: Estimates of Sym. CGARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
C(4)	*		*	*	
C(5)	*	*	*	*	*
C(6)	*	*	*	*	*
C(7)	***			*	
C(8)	(+0.343) *		(-0.039) ***	(+0.253) *	
C(9)	*	**	*	*	*
C(7)+C(9)					

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(4) - p-values of the constant coefficient in the long run component of CGARCH(1,1);
C(5) - p-values of powers in the long run component of CGARCH(1,1);
C(6) - p-values of the coefficient of the difference of (ARCH term - GARCH term) in the long run component of CGARCH(1,1);
C(7) - p-values of the ARCH coefficient in the transitory component of CGARCH(1,1);
C(8) - p-values of the asymmetry coefficient in the transitory component of CGARCH(1,1);
C(9) - p-values of the GARCH coefficient in the transitory component of CGARCH(1,1).
(+/-) - the positive / negative coefficient value.

Table B44b: Estimates of Asym.CGARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
Lag 1					^ , "
Lag 2				^^^ , """	^^ , ""
Lag 3					^^ , ""
Lag 4					^^ , ""
Lag 5					^^^ , """
Lag 6					^^^ , """
Lag 7					^^^ , """
Lag 8					
Lag 9					

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B45a: ARCH effects in Sym. CGARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	ρ(π)=0	φ=0.995
Lag 1				^ , "	
Lag 2	^^^ , """			^ , "	
Lag 3				^ , "	
Lag 4				^ , "	
Lag 5				^ , "	
Lag 6				^ , "	
Lag 7				^ , "	
Lag 8				^ , "	
Lag 9		^^^ , """		^ , "	

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B45b: ARCH effects in Asym. CGARCH(1,1)

B.2.4 Residuals μ_t

Residuals μ_t are produced by Equation (B.4).

Table B46 reports ARCH tests for the residuals μ_t. Both the F-statistic and LM-statistic are highly significant (at 1%) at all lags (1 to 9) in the baseline and sensitivity analysis, which is same as that reported in Tables B1, B16, and B31. The innovations μ_t show the evidence of ARCH effects.

We estimate the GARCH type models and test additional ARCH up to order 9.

Symmetric conditional volatility models are reported in Tables B47 - B52.

Tables B47 and B48 report the estimates and tests of additional ARCH effects for the ARCH(1) model. All coefficients in the conditional variance equation are highly significant (at 1%) in the baseline and sensitivity analysis. ARCH effects are significant at all lags (1-9) in $\phi = 0.995$ while no additional ARCH is found in the baseline and $\rho_\pi = 0$, which is similar as that reported in Tables B18 and B33. Additional ARCH effects also exist at the lags 4-9 in $\gamma = 0.7$ and at the lags 3-9 in $\delta = -0.0025$.

We estimate the GARCH(1,1) model in Table B49 and test its additional ARCH up to order 9 in Table B50. In Table B49, the coefficients on all three terms in the conditional variance equation are highly significant (at 1%) for both the baseline and sensitivity analysis. There is no difference changing each parameter separately. After the estimation of the GARCH(1,1) model, additional ARCH effects exist in the residuals for the sensitivity analysis while no additional ARCH is found in the baseline. Specifically, the F-statistic and LM-statistic are significant at all lags (up to 9 lags) in $\gamma = 0.7$, $\delta = -0.0025$ and $\phi = 0.995$, which is same as those in Tables B5, B20, and B35, and at the lag 1 in $\rho_\pi = 0$.

The estimates and tests for the GARCH-M are reported in Tables B51 and B52. All constants, ARCH and GARCH terms in the conditional variance equation are highly significant (at 1%), which is the same as those in Tables B6, B21 and B36. GARCH-M has exactly same additional ARCH as GARCH(1,1) has in Table B50. In Table B52, additional ARCH effects exist at all lags (1-9) in $\gamma = 0.7$, $\delta = -0.0025$ and $\phi = 0.995$, which is the same as those found in Tables B7, B22 and B37, and at the lag 1 in $\rho_\pi = 0$, which is similar to those in Tables B7 and B22. We find, but not report, that the parameters of the conditional variance term in the mean equation are positive and highly significant (at 1%) in $\gamma = 0.7$ and $\rho_\pi = 0$. It suggests that the innovations $\widehat{\zeta}_t$ are partly determined by its risk.

Asymmetric conditional volatility models are reported in Tables B53 - B60.

In Table B53, the estimates for the EGARCH(1,1) model are reported. All of the constants, ARCH and GARCH coefficients are highly significant (at 1%) for both the baseline and sensitivity analysis, which is same as those reported in Tables B23 and B38. The parameters of the asymmetric terms are significant and positive (+0.032 at 5% in $\gamma = 0.7$and $+\ 0.039$ at 1% in $\rho_\pi = 0$) in the sensitivity analysis. In Table B54, the additional ARCH effects up to order 9 are present in the sensitivity analysis, which is the same as those in Table B24 and similar to those in Table B9. The baseline has no ARCH effects left, which is the same as those in Table B39 but different to those in Tables B9 and B24 where ARCH effects are present in the baseline.

The results for the TARCH model are reported in Tables B55 and B56. All of the constants, ARCH and GARCH coefficients are highly significant (at 1%), which is same as those reported in Tables B10, B25 and B40. The baseline, $\gamma = 0.7$ and $\rho_\pi = 0$ have asymmetric effects. The coefficients of the asymmetric terms are significant (at 5%, 1% and 1% in the baseline, $\gamma = 0.7$ and $\rho_\pi = 0$, respectively) with the positive (+0.347 in the baseline) and negative (-0.236 in $\gamma = 0.7$ and -1.314 in $\rho_\pi = 0$) values in terms of asymmetric news and leverage effect. It is common that in $\rho_\pi = 0$ the TARCH(1,1) model is able to capture asymmetric effects shown in Tables B10, B25, and B55 while in the baseline it is also capable of capturing the asymmetric effects as displayed in Tables B40 and B55. In Table B56, it is found that TARCH(1,1) has same ARCH effects as the GARCH(1,1) has as in Table B50. The F-statistic and LM-statistic are highly significant (at 1%) at all lags in $\gamma = 0.7$, $\delta = -0.0025$ and $\phi = 0.995$, which is the same as those found in Tables B11, B26 and B41.

Tables B57 and B58 display the estimates and ARCH tests for the PARCH(1,1) model. In Table B57, all of the ARCH and GARCH coefficients and the power parameters of the standard deviation are highly significant (at 1%) in both the baseline and sensitivity analysis, which is the same as those in Tables B12 and B42. The constants in the sensitivity analysis are highly significant (at 1%) while the constant in the baseline is not significant. In the baseline, $\gamma = 0.7$, $\delta = -0.0025$ and $\rho_\pi = 0$, the asymmetric effects are present at the 10%, 1%, 1% and 1% significance level, respectively, with the positive (+0.077 in the baseline) and negative (-0.127 in $\gamma = 0.7$, -0.161 in $\delta = -0.0025$ and -0.205 in $\rho_\pi = 0$) values. It is common that in $\gamma = 0.7$ and $\rho_\pi = 0$ the PARCH(1,1) model captures the asymmetric effects shown in Tables B27 and B57. In Table B58, both the F-statistic and LM-statistic are significant at lags 8-9 in $\gamma = 0.7$, at all lags (1-9) in $\delta = -0.0025$ and at the most lags except at the lags 3 and 8 in $\phi = 0.995$. For the PARCH model, ARCH effects do not exist in the baseline while additional ARCH effects are present in the sensitivity analysis, which is the same as that reported in Tables B13, B28, and B43.

The estimates and additional ARCH tests for the symmetric CGARCH(1,1) are reported in Tables B59a and B60a. In the permanent equation, all constants are not significant while all coefficients of the powers are highly significant (at 1%) for both the baseline and sensitivity analysis. The parameter estimate on the difference term in $\rho_\pi = 0$ is not significant while the others are highly significant (at 1%). In the transitory equation, all ARCH terms are not significant while all GARCH terms in sensitivity analysis except the non-significant one in the baseline are highly significant (at 1%). Additional ARCH effects are found significantly (at 10% at the lag 1 and 2) in the baseline and (at 10% at the lag 2) in $\rho_\pi = 0$. Comparing Tables B15a, B30a, B45a and B60a, it is found that in common there is no additional ARCH in $\gamma = 0.7$ and $\delta = -0.0025$.

Tables B59b and B60b report the results of ARCH estimates and tests for the asymmetric CGARCH(1,1) model. In the permanent equation, all coefficients of the powers and the difference term are highly significant (at 1%). The constants and ARCH coefficients are significant in both the baseline and $\rho_\pi = 0$, which is also found in Tables B29b and B44b. The GARCH term in the transitory equation is highly significant (at 1%) in the baseline, $\rho_\pi = 0$ and $\phi = 0.995$. It is found that in common the GARCH term in the baseline is significant for Tables B14b, B29b, B44b and B59b. CGARCH(1,1) successfully captures asymmetric effects in the baseline, $\delta = -0.0025$, $\rho_\pi = 0$ and $\phi = 0.995$ at the 1%, 10%, 1% , and 5% significance levels respectively with the positive (+0.275), negative (-0.084), negative (-0.253), and positive (+0.042) coefficient values respectively. The asymmetric effects found in the baseline, $\delta = -0.0025$, $\rho_\pi = 0$ are common for Tables B29b, B44b and B59b. In Table B60b, the presence of additional ARCH is in $\rho_\pi = 0$ for most lags except at the lag 1. Comparing all tables of additional ARCH for the asymmetric CGARCH (Tables B15b, B30b, B45b and B60b), it is found that additional ARCH is present in the baseline in both Tables B30b and B45b, and exists in sensitivity analysis in Tables B45b and B60b. There is no additional ARCH found in $\delta = -0.0025$ and $\phi = 0.995$ as in all of Tables B15b, B30b, B45b and B60b for the asymmetric CGARCH(1,1) model.

Testing and Estimating ARCH for the residuals μ_t which are produced by Equation (B.4)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
Lag 1	^, "	^, "	^, "	^, "	^, "
Lag 2	^, "	^, "	^, "	^, "	^, "
Lag 3	^, "	^, "	^, "	^, "	^, "
Lag 4	^, "	^, "	^, "	^, "	^, "
Lag 5	^, "	^, "	^, "	^, "	^, "
Lag 6	^, "	^, "	^, "	^, "	^, "
Lag 7	^, "	^, "	^, "	^, "	^, "
Lag 8	^, "	^, "	^, "	^, "	^, "
Lag 9	^, "	^, "	^, "	^, "	^, "

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B46: ARCH effects in μ_t

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
C(2)	*	*	*	*	*
C(3)	*	*	*	*	*

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(2) - p-values of the constant coefficient in the variance equation of ARCH (1);
C(3) - p-values of the ARCH coefficient in the variance equation of ARCH (1).

Table B47: Estimates of ARCH(1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
Lag 1					^^^, """"
Lag 2					^, "
Lag 3			^^, """		^, "
Lag 4		^, "	^^, """		^, "
Lag 5		^, "	^, "		^, "
Lag 6		^, "	^, "		^, "
Lag 7		^, "	^, "		^, "
Lag 8		^, "	^, "		^, "
Lag 9		^, "	^, "		^, "

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B48: ARCH effects in ARCH(1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
C(2)	*	*	*	*	*
C(3)	*	*	*	*	*
C(4)	*	*	*	*	*
C(3)+C(4)					

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(2) - p-values of the constant coefficient in the variance equation of GARCH (1,1);
C(3) - p-values of the ARCH coefficient in the variance equation of GARCH (1,1);
C(4) - p-values of the GARCH coefficient in the variance equation of GARCH (1,1).

Table B49: Estimates of GARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
Lag 1		^, "	^, "	^^^, """"	^, "
Lag 2		^, "	^, "		^, "
Lag 3		^, "	^, "		^, "
Lag 4		^, "	^, "		^, "
Lag 5		^, "	^, "		^, "
Lag 6		^, "	^, "		^, "
Lag 7		^, "	^, "		^, "
Lag 8		^, "	^, "		^, "
Lag 9		^, "	^, "		^, "

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B50: ARCH effects in GARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
C(3)	*	*	*	*	*
C(4)	*	*	*	*	*
C(5)	*	*	*	*	*
C(4)+C(5)					

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(3) - p-values of the constant coefficient in the variance equation of GARCH-M;
C(4) - p-values of the ARCH coefficient in the variance equation of GARCH-M;
C(5) - p-values of the GARCH coefficient in the variance equation of GARCH-M.

Table B51: Estimates of GARCH-M

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
Lag 1		^, "	^, "	^^^, """	^, "
Lag 2		^, "	^, "		^, "
Lag 3		^, "	^, "		^, "
Lag 4		^, "	^, "		^, "
Lag 5		^, "	^, "		^, "
Lag 6		^, "	^, "		^, "
Lag 7		^, "	^, "		^, "
Lag 8		^, "	^, "		^, "
Lag 9		^, "	^, "		^, "

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B52: ARCH effects in GARCH-M

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
C(2)	*	*	*	*	*
C(3)	*	*	*	*	*
C(4)		(+0.032) **		(+0.039) *	
C(5)	*	*	*	*	*
C(3)+C(5)					

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(2) - p-values of the constant coefficient in the variance equation of EGARCH (1,1);
C(3) - p-values of the ARCH coefficient in the variance equation of EGARCH (1,1);
C(4) - p-values of the asymmetry term in the variance equation of EGARCH (1,1);
C(5) - p-values of the GARCH coefficient in the variance equation of EGARCH (1,1).
(+) - the positive coefficient value.

Table B53: Estimates of EGARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
Lag 1		^^, ""	^^^, """	^, "	^^, ""
Lag 2		^^, ""	^^, ""	^, "	^, "
Lag 3		^, "	^^, ""	^, "	^, "
Lag 4		^, "	^, "	^, "	^, "
Lag 5		^, "	^, "	^, "	^, "
Lag 6		^, "	^, "	^, "	^, "
Lag 7		^, "	^, "	^, "	^, "
Lag 8		^, "	^, "	^, "	^, "
Lag 9		^, "	^, "	^, "	^, "

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B54: ARCH effects in EGARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
C(2)	*	*	*	*	*
C(3)	*	*	*	*	*
C(4)	(+0.347) **	(-0.236) *		(-1.314) *	
C(5)	*	*	*	*	*
C(3)+C(5)					

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(2) - p-values of the constant coefficient in the variance equation of TARCH(1,1);
C(3) - p-values of the ARCH coefficient in the variance equation of TARCH(1,1);
C(4) - p-values of the threshold term in the variance equation of TARCH(1,1);
C(5) - p-values of the GARCH coefficient in the variance equation of TARCH(1,1).
(+/-) - the positive / negative coefficient value.

Table B55: Estimates of TARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
Lag 1		^, "	^, "	^^^, """	^, "
Lag 2		^, "	^, "		^, "
Lag 3		^, "	^, "		^, "
Lag 4		^, "	^, "		^, "
Lag 5		^, "	^, "		^, "
Lag 6		^, "	^, "		^, "
Lag 7		^, "	^, "		^, "
Lag 8		^, "	^, "		^, "
Lag 9		^, "	^, "		^, "

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B56: ARCH effects in TARCH(1,1)

	Baseline	γ=0.7	δ=-0.0025	$\rho(\pi)$=0	φ=0.995
C(2)		*	*	*	*
C(3)	*	*	*	*	*
C(4)	(+0.077) ***	(-0.127) *	(-0.161) *	(-0.205) *	
C(5)	*	*	*	*	*
C(6)	*	*	*	*	*
C(3)+C(5)					

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(2) - p-values of the constant coefficient in the std dev equation of PARCH(1,1);
C(3) - p-values of the ARCH coefficient in the std dev equation of PARCH(1,1);
C(4) - p-values of the asymmetry effects in the std dev equation of PARCH(1,1);
C(5) - p-values of the GARCH coefficient in the std dev equation of PARCH(1,1).
C(6) - p-values of the power parameter of the standard deviation in PARCH(1,1);
(+/-) - the positive / negative coefficient value.

Table B57: Estimates of PARCH(1,1)

	Baseline	*γ=0.7*	*δ=-0.0025*	*ρ(π)=0*	*φ=0.995*
Lag 1			^^^, """		^^, ""
Lag 2			^^^, """		^^^, """
Lag 3			^, "		
Lag 4			^, "		^^^, """
Lag 5			^^, ""		^^^, """
Lag 6			^^, ""		^^^, """
Lag 7			^^, ""		^^^, """
Lag 8		^^^, """	^^, ""		
Lag 9		^^, ""	^^^, """		^^^, """

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B58: ARCH effects in PARCH(1,1)

	Baseline	*γ=0.7*	*δ=-0.0025*	*ρ(π)=0*	*φ=0.995*
C(2)					
C(3)	*	*	*	*	*
C(4)	*	*	*		*
C(5)					
C(6)		*	*	*	*
C(5)+C(6)					

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(2) - p-values of the constant coefficient in the long run component of CGARCH(1,1);
C(3) - p-values of powers in the long run component of CGARCH(1,1);
C(4) - p-values of the coefficient of the difference of (ARCH term - GARCH term) in the long run component of CGARCH(1,1);
C(5) - p-values of the ARCH coefficient in the transitory component of CGARCH(1,1);
C(6) - p-values of the GARCH coefficient in the transitory component of CGARCH(1,1).

Table B59a: Estimates of Sym. CGARCH(1,1)

	Baseline	*γ=0.7*	*δ=-0.0025*	*ρ(π)=0*	*φ=0.995*
C(2)	*			*	
C(3)	*	*	*	*	*
C(4)	*	*	*	*	*
C(5)	*			*	
C(6)	(+0.275) *		(-0.084) ***	(-0.253) *	(+0.042) **
C(7)	*			*	*
C(5)+C(7)					

At *(1%), **(5%), ***(10%) significant levels for z-statistics;
C(2) - p-values of the constant coefficient in the long run component of CGARCH(1,1);
C(3) - p-values of powers in the long run component of CGARCH(1,1);
C(4) - p-values of the coefficient of the difference of (ARCH term - GARCH term) in the long run component of CGARCH(1,1);
C(5) - p-values of the ARCH coefficient in the transitory component of CGARCH(1,1);
C(6) - p-values of the asymmetry coefficient in the transitory component of CGARCH(1,1);
C(7) - p-values of the GARCH coefficient in the transitory component of CGARCH(1,1);
(+/-) - the positive / negative coefficient value.

Table B59b: Estimates of Asym.CGARCH(1,1)

	Baseline	*γ=0.7*	*δ=-0.0025*	*ρ(π)=0*	*φ=0.995*
Lag 1	^^^, """				
Lag 2	^^^, """			^^^, """	
Lag 3					
Lag 4					
Lag 5					
Lag 6					
Lag 7					
Lag 8					
Lag 9					

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B60a: ARCH effects in Sym. CGARCH(1,1)

	Baseline	*γ=0.7*	*δ=-0.0025*	*ρ(π)=0*	*φ=0.995*
Lag 1					
Lag 2				^, "	
Lag 3				^, "	
Lag 4				^, "	
Lag 5				^, "	
Lag 6				^, "	
Lag 7				^, "	
Lag 8				^, "	
Lag 9				^, "	

At ^(1%), ^^(5%), ^^^(10%) significant levels for F-statistics;
At "(1%), ""(5%), """(10%) significant levels for Chi-square statistics.

Table B60b: ARCH effects in Asym. CGARCH(1,1)

B.2.5 Summary

As stated above, we test and estimate ARCH in both the baseline and sensitivity analysis for the residuals τ_t (ϱ_t, ε_t, ξ_t and μ_t) produced by the four different conditional mean equations (Equations B.1, B.2, B.3, and B.4). We report the results in terms of the significant coefficients, asymmetric effects, persistent shocks, and the relationship of return with risk. We find that the results for Equations (B.3) and (B.4) are not only similar to but also have even more information than those found in Equations (B.1) and (B.2). For example, the GARCH-M model shows that Moore and Roche's model is able to capture the relationship between the return and risk where the return is partly determined by its risk while the CGARCH is capable of capturing additional ARCH. We find the best fit GARCH models.

The findings are summarized as follows:

1 ARCH effects: the presence of ARCH effects is in all innovations τ_t.

2 Symmetric conditional volatility models (ARCH, GARCH, GARCH-M):

a Coefficients:

i The coefficients on all terms in the conditional variance equation are significant.

ii Neither of the performance in the baseline and sensitivity analysis dominates.

b ARCH effects:

i GARCH(1,1) is the best fit symmetric model for capturing additional ARCH. (*models criteria*)

(GARCH-M has the almost same ARCH effects as GARCH(1,1) does; ARCH(1,1) ranks 3rd.)

ii $\phi = 0.995$ is the best in the sensitivity analysis for capturing additional ARCH. (*sensitivity analysis criteria*)

($\gamma = 0.7$ and $\delta = -0.0025$ rank 2nd; $\rho_{\pi} = 0$ ranks 4th; the baseline ranks 5th.)

iii The sensitivity analysis dominates.

c Returns and risk: the filtered spot returns Δs_t^f and innovations $\widehat{\zeta}_t$ are partly determined by their own risk, respectively. Using the *GARCH-M* models the conditional variance term that appears in the mean equation is positive and significant (at 1% or 5%) in $\gamma = 0.7$ and $\rho_{\pi} = 0$, respectively. It suggests that the model is capable of capturing the relationship of return with risk: higher market risk (proxied by the conditional variance) will lead to higher returns.

3 Asymmetric conditional volatility models (*EGARCH, GJR/TARCH, PARCH, CGARCH)*:

a Coefficients:

i The significance levels of the coefficients on the terms in the conditional variance equation change. Most of them are significant.

ii Neither of the performances in the baseline and sensitivity analysis dominates.

b ARCH effects:

i EGARCH(1,1) is the best fit model for capturing additional ARCH. (*models criteria*)

(TARCH ranks 2nd; PARCH ranks 3rd; asymmetric CGARCH ranks 4th; symmetric CGARCH ranks 5th.)

ii $\gamma = 0.7$ is the best in the sensitivity analysis for capturing additional ARCH. (*sensitivity analysis criteria*)

($\delta = -0.0025$ ranks 2nd; $\phi = 0.995$ ranks 3rd; $\rho_{\pi} = 0$ ranks 4th; the baseline ranks 5th.)

iii The sensitivity analysis dominates.

c Asymmetric effects:

i The asymmetric CGARCH(1,1) is the best fit model for capturing asymmetric effects. (*models criteria*)

(EGARCH ranks 2nd; TARCH and PARCH rank 3rd; CGARCH ranks 5th.)

ii $\rho_\pi = 0$ is the best in the sensitivity analysis for capturing asymmetric effects. (*sensitivity analysis criteria*)

(The baseline ranks 2nd; $\phi = 0.995$ and $\gamma = 0.7$ rank 3rd; $\delta = -0.0025$ ranks 5th.)

iii The sensitivity analysis dominates.

d Persistent shocks:

i PARCH(1,1) is the best fit model for capturing highly persistent shocks to the conditional variance. (*models criteria*)

(CGARCH ranks 2nd; other asymmetric conditional volatility models are unable to explore this property.)

ii $\gamma = 0.7$ is the best in the sensitivity analysis for capturing highly persistent shocks to the conditional variance. (*sensitivity analysis criteria*)

($\rho_\pi = 0$ ranks 2nd; $\phi = 0.995$ and $\delta = -0.0025$ rank 3rd; the baseline ranks 5th.)

iii The sensitivity analysis dominates.

e Both asymmetric effects and persistent volatility shocks:

i PARCH(1,1) is the best fit model for capturing both asymmetric effects and persistent volatility shocks. (*models criteria*)

(CGARCH ranks 2nd; other asymmetric conditional volatility models are unable to explore both properties.)

ii $\rho_\pi = 0$ is the best in the sensitivity analysis for capturing both asymmetric effects and persistent volatility shocks. (*sensitivity analysis criteria*)

($\gamma = 0.7$ ranks 2nd; $\phi = 0.995$ rank 3rd; the baseline ranks 4th; $\delta = -0.0025$ rank 5th.)

iii The sensitivity analysis dominates.

4 The baseline and sensitivity analysis: the sensitivity analysis dominates.

Basing on the analysis and summary for the ARCH effects and estimates for the theoretical residuals in baseline and sensitivity analysis, the main conclusions are:

- GARCH class conditional volatility models are appropriate for the theoretical (simulated) quarterly FOREX data.
- The predictable properties of conditional volatility are found in the theoretical (simulated) quarterly FOREX data.
- Additional ARCH is significant after the standard GARCH class models' estimates.
- In symmetric conditional volatility models, the GARCH(1,1) model is the best fit model to conditional volatility of innovations.
- In asymmetric conditional volatility models, for conditional volatility of innovations, the EGARCH(1,1) model is the best fit model to additional ARCH; the CGARCH model (we refer the asymmetric CGARCH model as CGARCH for short afterwards) is the best fit model for asymmetric effects; the PARCH(1,1) model is the best fit model for highly persistent volatility shocks.
- Sensitivity analysis is superior to the baseline in terms of ARCH effects, asymmetric effects and persistent volatility shocks. $\gamma = 0.7$ is recommended due to the

estimating and additional ARCH performances for the symmetric and asymmetric conditional volatility models.

- The theoretical model (Moore and Roche's model) can generate the required conditional volatility, even the asymmetric conditional volatility.
- The theoretical model (Moore and Roche's model) can capture the relationship of returns with risk.

In a word, the ARCH estimates and tests, as well as steps in model selection process for the simulated data, are reported in detail. The results suggest that the best fitting GARCH class models are CGARCH(1,1), EGARCH(1,1) and PARCH(1,1).

Appendix C.1

C.1 Quadratic utility function

$$\max_{\alpha_t} E_t(U_{t+1}) \equiv E_t(w_{t+1} - 0.5\gamma w_{t+1}^2)$$

subject to

$$w_{t+1} = \alpha_t r_{t+1} + (1 - \alpha_t) r_{f,t+1}$$

Differentiate with respect to α_t:

$$\begin{aligned} \frac{d}{d\alpha_t} U(w_{t+1}) &= U'(w_{t+1}) \frac{dw_{t+1}}{d\alpha_t} \\ &= (r_{t+1} - r_{f,t+1})\{1 - \gamma [\alpha_t (r_{t+1} - r_{f,t+1}) + r_{f,t+1}]\} \end{aligned}$$

Take expected values, set equal to zero, and solve for α_t

$$\alpha_t = E_t \left[\frac{1 - \gamma r_{f,t+1}}{\gamma \varrho_{t+1}} \right]$$

where $\varrho_{t+1} = r_{t+1} - r_{f,t+1}$. Multiplying both the denominator and numerator by ϱ_{t+1} gives

$$\begin{aligned} \alpha_t &= \frac{(1 - \gamma r_{f,t+1}) E_t(\varrho_{t+1})}{\gamma \left[E_t(\varrho_{t+1}^2) - (E_t(\varrho_{t+1}))^2 + (E_t(\varrho_{t+1}))^2 \right]} \\ &= \frac{\mu_{t+1}(1 - \gamma r_{f,t+1})}{\gamma \left(\mu_{t+1}^2 + \sigma_{t+1}^2 \right)} \end{aligned}$$

where $\mu_{t+1} \equiv E(\varrho_{t+1})$ and σ_{t+1}^2 is the true conditional variance of the conditional expected asset return ϱ_{t+1}. Investors in this model are assumed not to know the true parameters of the model and must try and find an estimate of $\widehat{\sigma}_{t+1}^2$ in order to work out their best guess of optimal α_t. Therefore, the optimal portfolio weight of the risky asset as follows

$$\alpha_t = \frac{\mu_{t+1}(1 - \gamma r_{f,t+1})}{\gamma(\mu_{t+1}^2 + \widehat{\sigma}_{t+1}^2)}$$

where $\widehat{\sigma}_{t+1}^2$ is the estimated conditional variance of σ_{t+1}^2 from the forecasting models.

Wealth at time $t+1$ becomes

$$\begin{aligned} w_{t+1} &= \left[\frac{\mu_{t+1}(1-\gamma r_{f,t+1})}{\gamma(\mu_{t+1}^2+\widehat{\sigma}_{t+1}^2)}\right] r_{t+1} + \left[1-\frac{\mu_{t+1}(1-\gamma r_{f,t+1})}{\gamma(\mu_{t+1}^2+\widehat{\sigma}_{t+1}^2)}\right] r_{f,t+1} \\ &= \frac{\mu_{t+1}(1-\gamma r_{f,t+1})\varrho_{t+1}}{\gamma(\mu_{t+1}^2+\widehat{\sigma}_{t+1}^2)} + r_{f,t+1} \end{aligned}$$

Investors' expected utility may then be written

$$\begin{aligned} & E_t(\widehat{U}_{t+1}) \\ =\ & E_t(w_{t+1} - 0.5\gamma w_{t+1}^2) \\ =\ & E_t\{\frac{\mu_{t+1}(1-\gamma r_{f,t+1})\varrho_{t+1}}{\gamma(\mu_{t+1}^2+\widehat{\sigma}_{t+1}^2)} + r_{f,t+1} - 0.5\frac{\mu_{t+1}^2(1-\gamma r_{f,t+1})^2\varrho_{t+1}^2}{\gamma(\mu_{t+1}^2+\widehat{\sigma}_{t+1}^2)^2} \\ & -\frac{\mu_{t+1}(1-\gamma r_{f,t+1})\varrho_{t+1}r_{t+1}}{\mu_{t+1}^2+\widehat{\sigma}_{t+1}^2} - 0.5\gamma r_{t+1}^2\} \\ =\ & E_t\left\{r_{f,t+1} - 0.5\gamma r_{t+1}^2 + \frac{\mu_{t+1}^2(1-\gamma r_{f,t+1})^2}{\gamma}\left[\frac{E_t\left(\varrho_{t+1}\right)}{\mu_{t+1}(\mu_{t+1}^2+\widehat{\sigma}_{t+1}^2)} - 0.5\frac{E_t(\varrho_{t+1}^2)}{(\mu_{t+1}^2+\widehat{\sigma}_{t+1}^2)^2}\right]\right\} \\ =\ & r_{f,t+1} - 0.5\gamma r_{t+1}^2 + \frac{\mu_{t+1}^2(1-\gamma r_{f,t+1})^2}{\gamma}\left[\frac{1}{\mu_{t+1}^2+\widehat{\sigma}_{t+1}^2} - 0.5\frac{E_t(\varrho_{t+1}^2)-(E_t(\varrho_{t+1}))^2+(E_t(\varrho_{t+1}))^2}{(\mu_{t+1}^2+\widehat{\sigma}_{t+1}^2)^2}\right] \\ =\ & r_{f,t+1} - 0.5\gamma r_{t+1}^2 + \frac{\mu_{t+1}^2(1-\gamma r_{f,t+1})^2}{\gamma}\left[\frac{1}{\mu_{t+1}^2+\widehat{\sigma}_{t+1}^2} - 0.5\frac{\sigma_{t+1}^2+\mu_{t+1}^2}{(\mu_{t+1}^2+\widehat{\sigma}_{t+1}^2)^2}\right] \end{aligned}$$

In the chapter, using squared excess return ϱ_{t+1}^2 as the proxy for daily volatility gives

$$E_t(\widehat{U}_{t+1}) = r_{f,t+1} - 0.5\gamma r_{t+1}^2 + \frac{\mu_{t+1}^2(1-\gamma r_{f,t+1})^2}{\gamma}\left[\frac{1}{\mu_{t+1}^2+\widehat{\sigma}_{t+1}^2} - 0.5\frac{\varrho_{t+1}^2+\mu_{t+1}^2}{(\mu_{t+1}^2+\widehat{\sigma}_{t+1}^2)^2}\right]$$

Appendix C.2

C.2 Exponential utility function

$$\max_{g_t} E_t[u\,(w_{t+1})] = -e^{-A\mu_{w_{t+1}}+(A^2/2)\sigma_{w_{t+1}}^2}, \quad A>0$$

subject to

$$w_{t+1} = g_t r_{t+1} + (1-g_t) r_{f,t+1}$$

where $\mu_{w_{t+1}} = E_t(w_{t+1})$ and $\sigma^2_{w_{t+1}} = E_t\left[(w_{t+1} - \mu_{w_{t+1}})^2\right]$ denote the expected return and variance of the portfolio. Plugging the budget constraint into the expected exponential utility function gives

$$\begin{aligned}
& \max_{g_t} E_t[u\,(w_{t+1})] \\
= \;& -e^{-A[g_t E_t(r_{t+1})+(1-g_t)r_{f,t+1}]+(A^2/2)[E_t(w^2_{t+1})-(E_t(w_{t+1}))^2]} \\
= \;& -e^{-A[g_t E_t(r_{t+1})+(1-g_t)r_{f,t+1}]+(A^2/2)[E_t(g_t r_{t+1}+(1-g_t)r_{f,t+1})^2-\left(E_t\left(g_t r_{t+1}+(1-g_t)r_{f,t+1}\right)\right)^2]} \\
= \;& -e^{-A\{g_t E_t(r_{t+1})+(1-g_t)r_{f,t+1}-(A/2)[E_t(g_t^2 r^2_{t+1}+2g_t r_{t+1}(1-g_t)r_{f,t+1}+(1-g_t)^2 r^2_{f,t+1})-[g_t E_t(r_{t+1})+(1-g_t)r_{f,t+1}]^2\}} \\
= \;& -e^{-A[g_t E_t(r_{t+1})+(1-g_t)r_{f,t+1}]+(A^2/2)g_t^2\{E_t(r^2_{t+1})-[E_t(r_{t+1})]^2\}}
\end{aligned}$$

Differentiate with respect to g_t, set equal to zero, and solve for g_t

$$\begin{aligned}
& \frac{d}{dg_t}u\,(w_{t+1}) \\
= \;& u'(w_{t+1})\,\frac{dw_{t+1}}{dg_t} \\
= \;& -e^{-A[g_t E_t(r_{t+1})+(1-g_t)r_{f,t+1}]+(A^2/2)g_t^2\{E_t(r^2_{t+1})-[E_t(r_{t+1})]^2\}}\{-A[E_t(r_{t+1}-r_{f,t+1})]+A^2 g_t \sigma^2_{r,t+1}\} \\
= \;& -e^{-A[g_t E_t(r_{t+1})+(1-g_t)r_{f,t+1}]+(A^2/2)g_t^2\{E_t(r^2_{t+1})-[E_t(r_{t+1})]^2\}}(-A\mu_{t+1}+A^2 g_t \sigma^2_{r,t+1}) \\
= \;& 0
\end{aligned}$$

The optimal fraction of the risky asset in the portfolio is

$$g_t = \frac{\mu_{t+1}}{A\widehat{\sigma}^2_{r,t+1}} = \frac{\mu_{t+1}}{A\widehat{\sigma}^2_{t+1}}$$

After substituting back in for optimal g_t, the maximized exponential function is

$$\begin{aligned}
E_t(\widehat{u}_{t+1}) &= -e^{-\mu^2_{t+1}/\widehat{\sigma}^2_{t+1}-Ar_{f,t+1}+(A^2/2)(\frac{\mu_{t+1}}{A\widehat{\sigma}^2_{t+1}})^2\sigma^2_{t+1}} \\
&= -e^{-\mu^2_{t+1}/\widehat{\sigma}^2_{t+1}-Ar_{f,t+1}+\frac{1}{2}\sigma^2_{t+1}\frac{\mu^2_{t+1}}{(\widehat{\sigma}^2_{t+1})^2}}
\end{aligned}$$

Bibliography

ABEL, A. (1990): "Asset prices under habit formation and catching-up-with-the-Joneses," *American Economic Review*, 80, 38–42.

AIVAZIAN, V., J. CALLEN, I. KRINSKY, AND C. KWAN (1983): "Mean-variance utility functions and the demand for risky assets: an empirical analysis using exible functional forms," *Journal of Financial and Quantitative Analysis*, 18, 411–424.

AKGIRAY, V. (1989): "Conditional heteroscedasticity in time series of stock returns: evidence and forecasts," *Journal of Business*, 62, 55–80.

ALEXANDER, C. (2001): *Market models*. Wiley, Chichester.

ALFORD, A., AND R. JAMES (1995): "Predicting long-term stock return volatility: implications for accounting and valuation of equity derivatives," *Accounting Review*, 70, 599–618.

ALIZADEH, S., M. BRANDT, AND F. DIEBOLD (1999): "Range-based estimation of stochastic volatility models or exchange rate dynamics are more interesting than you think," *Working paper, University of Pennsylvania.*

AMIN, K., AND V. NG (1997): "Inferring future volatility from the information in implied volatility in Eurodollar options: a new approach," *Review of Financial Studies*, 10, 333–367.

ANDERSON, T., AND T. BOLLERSLEV (1998): "Answering the skeptics: yes, standard volatility models do provide accurate forecasts," *International Economic Review*, 39, 885–905.

ANDERSON, T., T. BOLLERSLEV, P. CHRISTOFFERSEN, AND F. DIEBOLD (2005): *Volatility and correlation forecasting*, vol. 1. North Holland, Amsterdam.

ANDERSON, T. G.AND BOLLERSLEY, T., F. X. DIEBOLD, AND P. LABYS (2000): "The distribution of realized exchange rate volatility," *PIER Working paper, North Western University, Duke University and University of Pennsylvania.*

ANDERSON, T. G.AND BOLLERSLEY, T., AND S. LANGE (1999): "Forecasting financial market volatility: sample frequency vis--vis forecast horizon," *Journal of Empirical Finance*, 6, 457–477.

Anderson, T. G., T. Bollersley, and F. X. Diebold (2003): "Some like it smooth, and some like it rough: untangling continuous and jump components in measuring, modelling, and forecasting asset return volatility," *CFS Working Paper No. 2003/35, SSRN.*

Artis, M., and W. Zhang (1997): "Volatility clustering and volatility transmission: a non-parametetric view of ERM exchange rates," *CEPR Discussion Paper*, p. No. 1594.

Aydemir, A. (1998): "Volatility modelling in finance," *in Forecasting Volatility in the Financial Markets, John Knight and Stephen Satchell, eds.*, Chapter 1, 1–46.

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

Baillie, R., A. Cecen, C. Erkal, and Y. Han (2004): "Measuring non-linearity, long memory and self-similarity in high-frequency European exchange rates," *Journal of International Financial Markets, Institutions and Money*, 14, 401–418.

Baillie, R., and M. King (1996): "Fractional differencing and long memory processes," *Journal of Econometrics*, 73, no.1, 1–324.

Balaban, E. (1999): "Forecasting stock market volatility: evidence from Turkey," *Unpublished Manuscript, Central Bank of the Republic of Turkey, and J.W. Goethe University, Frankfurt/Main, Germany.*

——— (2004): "Comparative forecasting performance of symmetric and asymmetric conditional volatility models of an exchange rate," *Economics Letters*, 83, 99–105.

——— (2005): "Forecasting stock market volatility: evidence from fourteen countries," *Working paper, University of Edinburgh, UK.*

Bank of England, U. K. (2008): "Annual Report 2008," *Main Publications, http://www.bankofengland.co.uk/publications/annualreport/2008/2008full.pdf.*

Barberis, N., M. Huang, and T. Santos (2001): "Prospect theory and asset prices," *Quarterly Journal of Economics*, 116, 1–53.

Barucci, E., and R. Reno (2002): "On measuring volatility and the GARCH forecasting performance," *Journal of International Financial Markets*, 12, 183–200.

Bauwens, L., W. Omrane, and P. Giot (2005): "News announcements, market activity and volatility in the euro/dollar foreign exchange market," *Journal of International Money and Finance*, 24, 1108–1125.

Bazdresch, S., and A. Werner (2005): "Regime switching models for the Mexican peso," *Journal of International Economics*, 65, 185–201.

Benavides, G. (2004): "Predictive accuracy of futures options implied volatility: the case of exchange rate futures Mexican Peso-U.S. Dollar," *Central Bank of Mexico*, working paper.

Bera, A., and M. Higgins (1993): "ARCH models: properties, estimation, and testing," *Journal of Economic Surveys*, 7, 305–365.

——— (1995): "On ARCH models: properties, estimating and testing," *in Surveys in Econometrics, L. Exley, D.A.R. George, C.J. Roberts and S. Sawyer, eds., reprinted from Journal of Economic Surveys.*

——— (1997): "ARCH and bilinearity as competing models for nonlinear dependence," *Journal of Business and Economic Statistics*, 15, 43–50.

BHATTACHARYYA, D. (1979): "On the validity of the quadratic utility approach in mean-variance portfolio analysis: an empirical test," *De Economist*, 127, 422–445.

BITTLINGMAYER, C. (1998): "Output, stock volatility, and political uncertainty in a natural experiment: Germany, 18801940," *Journal of Finance*, 53, 2243–2257.

BLACK, F. (1976): "Studies in stock price volatility changes," *Proceedings of the 1976 Business Meeting of the Business and Economics Statistics Section, American Statistical Association*, pp. 177–181.

BLACK, F., AND M. SCHOLES (1973): "The pricing of options and corporate liabilities," *Journal of Political Economy*, 31, 307–327.

BOLDRIN, M.AND CHRISTIANO, L., AND J. FISHER (2001): "Habit persistence, asset returns, and the business cycle," *American Economic Review*, 91, 149–166.

BOLLERSLEV, T. (1986): "Generalized autoregressive conditional heteroscedasticity," *Journal of Econometrics*, 31, 307–327.

BOLLERSLEV, T., Y. CHOU, AND K. KRONER (1992): "ARCH modelling in finance: a review of the theory and empirical evidence," *Journal of Econometrics*, 52, 5–59.

BOLLERSLEV, T., R. ENGLE, AND D. NELSON (1994): *ARCH models*, vol. 4. North Holland, Amsterdam.

BOLLERSLEV, T., AND E. GHYSELS (1996): "Periodic autoregressive conditional heteroscedasticity," *Journal of Business and Economic Statistics*, 14, 139–151.

BOLLERSLEV, T., AND D. JUBINSKI (1999): "Equity trading volume and volatility: latent information arrivals and common long-run dependencies," *Journal of Business and Economic Statistics*, 17, 9–21.

BOWERMAN, B., AND R. O'CONNELL (1979): *Forecasting and time series*. Duxbury Press, Massachusetts.

BOX, G., AND G. JENKINS (1970): *Time series analysis: forecasting and control.* Holder Day, San Francisco.

——— (1976): *Time series analysis: forecasting and control (revised ed.)*. Holder Day, San Francisco.

BRAILSFORD, T., AND R. FAFF (1996): "An evaluation of volatility forecasting techniques," *Journal of Banking and Finance*, 20, 419–438.

BREEDEN, D. (1979): "An intertemporal asset pricing model with stochastic consumption and investment opportunities," *Journal of Financial Economics*, 7, 265–296.

BROOKS, C. (1997): "Linear and nonlinear (non-) forecastability of daily sterling exchange rates," *Journal of Forecasting*, 16, 125–145.

BROOKS, C. (2002): *Introductory econometrics for finance*. Cambridge University Press, Cambridge.

BROOKS, C., AND S. BURKE (1998): "Forecasting exchange rate volatility using conditional variance models selected by information criteria," *Economics Letters*, 61, 273–278.

BROOKS, C., AND G. PERSAND (2003): "Volatility forecasting for risk management," *Journal of Forecasting*, 22, 1–22.

BROWN, S. (1990): *Estimating volatility*. Irwin, Chicago.

BRUNI, F., D. FAIR, AND R. O'BRIEN (1996): *Risk management in volatile financial markets*. Kluwer Academic Publishers Group, Netherlands.

CALVET, L.AND FISHER, A. (2001): "Forecasting multifractal volatility," *Journal of Econometrics*, 105, 27–58.

CALVET, L., A. FISHER, AND S. THOMPSON (2006): "Volatility comovement: a multifrequency approach," *Journal of Econometrics*, 131, 179–215.

CAMPA, J., AND P. CHANG (1998): "The forecasting ability of correlations implied in foreign exchange options," *Journal of International Money and Finance*, 17, 855–880.

CAMPBELL, J., AND J. COCHRANE (1999): "By Force of Habit: a Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, 205–251.

CAMPBELL, J., AND L. HENTSCHEL (1992): "No news is good news: an asymmetric model of changing volatility in stock returns," *Journal of Financial Studies*, 15, 615–648.

CAMPBELL, J., AND A. KYLE (1993): "Smart money, noise trading and stock price behaviour," *Review of Economic Studies*, 60, 1–34.

CAMPBELL, J., A. LO, AND A. MACKINLAY (1997): *The econometrics of financial markets*. Princeton University Press.

CECEN, A. A., AND C. ERKAL (1996): "Distinguishing between stochastic and deterministic behavior in high frequency foreign exchange rate returns: can non-linear dynamics help forecasting?," *International Journal of Forecasting*, 12, 465–473.

CHEN, A. S., AND M. T. LEUNG (1998): "Stochastic properties and predictability of intraday Taiwan exchange rates," *International Review of Financial Analysis*, pp. 207–220.

CHERNOV, M., AND E. GHYSELS (2000): "A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purposes of option valuation," *Journal of Financial Economics*, 56, 407–458.

CHIRAS, D., AND S. MANASTER (1978): "The information content of option prices and a test for market efficiency," *Journal of Financial Economics*, 50, 125–150.

CHO, D., AND E. FREES (1988): "Estimating the volatility of discrete stock prices," *Journal of Finance*, 43, 451–466.

Chopra, V., and W. Ziemba (1993): "The effect of errors in means, variances, and covariances on optimal portfolio choice," *Journal of Portfolio Management*, Winter, 6–11.

Christiano, L., M. Eichenbaum, and C. Evans (2005): "Nominal rigidities and the dynamic effects of a shock to monetary policy," *Journal of Political Economy*, 113, 1–45.

Christodoulakis, G., and S. Satchell (1998): "Hashing GARCH: a reassessment of volatility forecasting performance," *in Forecasting Volatility in the Financial Markets, John Knight and Stephen Satchell, eds.*, Chapter 6, 168–192.

Christoffersen, P., and F. Diebold (2000): "How relevant is volatility forecasting for financial risk management?," *Review of Economics and Statistics*, 82, 12–23.

Christoffersen, P., and K. Jacobs (2004): "The importance of the loss function in option valuation," *Journal of Financial Economics*, 72, 291–318.

Clements, M., and D. Hendry (1988): *Forecasting economic time series*. Cambridge University Press, New York.

——— (1999): *Forecasting non-stationary economic time series*. MIT Press, Cambridge, MA.

Cochrane, J. (2001): *Asset pricing*. Princeton University Press.

Cochrane, J., F. Longstaff, and P. Santa-Clara (2003): "Two trees: asset price dynamics induced by market clearing," *Working pape, Anderson Graduate School of Management, UCLA in its series, University of California at Los Angeles*, no. 1248.

Constantinides, G. (1990): "Habit formation: a resolution of the equity premium puzzle," *Journal of Political Economy*, 98, 519–543.

Cont, R. (2007): "Volatility clustering in financial markets: empirical facts and agent-based models," *Long Memory in Economics*, 2, 289–309.

Cornell, B. (1981): "The consumption based asset pricing model: a note on potential tests and applications," *Journal of Financial Economics*, 9, 103–108.

Cumby, R., S. Figlewski, and J. Hasbrouck (1993): "Forecasting volatilities and correlations with EGARCH models," *Journal of Derivatives*, 1, 51–63.

DeGennaro, R., and R. Shrieves (1997): "Public information releases, private information arrival, and volatility in the foreign exchange market," *Journal of Empirical Finance*, 4, 295–315.

Degiannakis, S., and E. Xekalaki (1999): "Predictability and model selection in the context of ARCH models," *Athens University of Economics and Business, Department of Statistics, Technical report*, 69.

——— (2007): "Assessing the performance of a prediction error criterion model selection algorithm in the context of ARCH models," *Applied Financial Economics*, 17, 149–171.

Demetrescu, M. (2007): "Optimal forecast intervals under asymmetric loss," *Journal of Forecasting*, 26, 227–238.

DIEBOLD, F. (2001): *Elements of forecasting*. Thomson Learning: South-Wester.

DIEBOLD, F., AND J. LOPEZ (1995): "Modelling volatility dynamics," *in Macroeconomics: Developments, Tensions and Prospects, Kevin D. Hoover, eds.*

——— (1996): "Forecast evaluation and combination," *in Handbook of Statistics, Maddala G.S. and Rao C.R., eds.*, 14, 214–268.

DIEBOLD, F., AND R. MARIANO (1995): "Comparing predictive accuracy," *Journal of Business and Economic Statistics*, 13, 252–263.

DIEBOLD, F., AND M. WATSON (1996): "New developments in economic forecasting," *special issue of Journal of Applied Econometrics*, 11, 453–594.

DIEBOLD, F. X. (1989): "Forecast combination and encompassing: reconciling two divergent literatures," *International Journal of Forecasting*, 5, 589–592.

DIMSON, E., AND P. MARSH (1990): "Volatility forecasting without data-snooping," *Journal of Banking of Finance*, 14, 399–421.

DING, Z., R. ENGLE, AND C. GRANGER (1993): "A long memory property of stock market returns and a new model," *Journal of Empirical Finance*, 1, 83–106.

DUESENBERRY, J. (1949): *Income, savings and the theory of consumer behavior*. Harvard University Press, Cambridge, MA.

DUNIS, C., J. LAWS, AND S. CHAUVIN (2000): "The use of market data and model combination to improve forecast accuracy," *Working paper, Liverpool Business School.*

——— (2002): "Forecasting and trading currency volatility: an application of recurrent neural regression and model combination," *Journal of Forecasting*, 21, 317–354.

EDERINGTON, L., AND W. GUAN (2002): "Forecasting volatility," *Working Paper, University of Oklahoma.*

EDEY, M., AND G. ELLIOT (1992): "Some evidence on option prices as predictors of volatility," *Oxford Bulletin of Economics and Statistics*, 54, 567–578.

EDMONDS, R. G., AND S. J. (2004): "Is exchange rate volatility excessive? An ARCH and AR approach," *The Quarterly Review of Economics and Finance*, 44, 122–154.

ELDER, A., AND G. GANNON (1998): "Evaluation of volatility forecasts in an economic value framework," *International Review of Financial Analysis*, 7, 221–236.

ENGLE, R. (1982): "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom in flation," *Econometrica*, 50, 987–1007.

——— (1993): "Statistical models for financial volatility," *Financial Analysts*, 49, 72–78.

ENGLE, R., AND T. BOLLERSLEV (1986): "Modelling the persistence of conditional variances," *Econometric Reviews*, 5, 1–50.

ENGLE, R., T. ITO, AND W. LIN (1990): "Meteor showers or heat waves? Heteroscedastic intradaily volatility in the foreign exchange market," *Econometrica*, 58, 525–542.

ENGLE, R., A. KANE, AND J. NOH (1997): "Index-option pricing with stochastic volatility and the value of accurate variance forecasts," *Review of Derivatives Research*, 1, 139–157.

ENGLE, R., AND G. LEE (1999): "A Permanent and transitory model of stock return volatility," *in R. Engle and H. White (eds.), Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W. J. Granger*, p. 475497.

ENGLE, R., AND J. MEZRICH (1997): "Fast acting asset allocation and tracking with time varying covariances," *Manuscript.*

ENGLE, R., AND V. NG (1993): "Measuring and testing the impact of news on volatility," *Journal of Finance*, 48, 1749–1778.

ENGLE, R., AND J. ROSENBERG (1995): "GARCH gamma," *Journal of Derivatives*, Winter, 47–59.

ENGLE, R. F., AND J. R. RUSSELL (1997): "Forecasting the frequency of changes in quoted foreign exchange prices with the autoregressive conditional duration model," *Journal of Empirical Finance*, 4, 187–212.

——— (1998): "Autoregressive conditional duration: a new model for irregularly spaced transaction data," *Econometrica*, 66, 1127–1162.

EVIEWS. (2007): *EViews 6 User's Guide.* Quantitative Micro Software, LLC, CA.

FAMA, E. (1963): "Mandelbrot and the stable paretian distribution," *Journal of Business*, 36, 420–429.

——— (1965): "The behavior of stock market prices," *Journal of Business*, 38, 34–105.

FAMA, E., AND K. FRENCH (1992): "The cross-section of expected stock returns," *Journal of Finance*, 47, 427–465.

——— (1993): "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics*, 33, 3–56.

——— (1996): "Multifactor explanations of asset pricing anomalies," *Journal of Finance*, 51, 55–84.

FIESS, N., AND R. MACDONALD (2002): "Towards the fundamentals of technical analysis: analyzing the information content of high, low and close prices," *Economic modelling*, 19, 353–374.

FIGLEWSKI, S. (1994): "Forecasting volatility using historical data," *New York University, Stern School of Business*, Discussion paper.

——— (1997): "Forecasting volatility, financial markets, institutions and instruments," *New York University Salomon Center*, 6, 1–88.

FIGLEWSKI, S., AND T. GREEN (1999): "Market risk and model risk for a

nancial institution writing options," *Journal of Finance*, 54, 1465–1999.

FLEMING, J., C. KIRBY, AND B. OSTDIEK (2001): "The economic value of volatility timing," *Journal of Finance*, 56, 329–352.

FRANSES, P., AND D. VAN DIJK (1996): "Forecasting stock market volatility using non-linear GARCH models," *Journal of Forecasting*, 15, 229–235.

FRENCH, K., G. SCHWERT, AND R. STAMBAUGH (1987): "Expected stock returns and volatility," *Journal of Financial Economics*, 19, 329.

FREUND, R. (1956): "The introduction of risk into a programming model," *Econometrica*, 24, 253–263.

FROMMEL, M. (2004): "Modelling exchange rate volatility in the run-up to EMU using a Markov Switching GARCH model," *University of Hannover, Discussion paper NO. 306, ISSN 0949-9962.*

FROMMEL, M., R. MACDONALD, AND L. MENKHOFF (2005): "Markov switching regimes in a monetary exchange rate model," *Economic Modelling*, 22, 485–502.

FUNG, H., C. LIE, AND A. MORENO (1990): "The forecasting performance of the implied standard deviation in currency options," *Managerial Finance*, 16, 24–29.

FUNG, W., AND D. HSIEH (1991): "Empirical Analysis of Implied Volatility: Stocks, Bonds and Currencies," *Working paper, Department of Finance, Fuqua School of Business.*

GALBRAITH, J., AND T. KISINBAY (2005): "Content horizons for conditional variance forecasts," *International Journal of Forecasting*, 21, 249–260.

GALLANT, A., P. ROSSI, AND G. TAUCHEN (1992): "Stock prices and volume," *Review of Financial Studies*, 5, 199–242.

GAVRISHCHAKA, V. V., AND S. B. GANGULI (2003): "Volatility forecasting from multiscale and high-dimensional market data," *Neurocomputing*, 55, 285–305.

GERLACH, S., S. RAMASWAMY, AND M. SCATIGNA (2006): "150 years of financial market volatility," *BIS Quarterly Review*, September, 77–91.

GERLOW, M., S. IRWIN, AND T. LIU (1993): "Economic evaluation of commodity price forecasting models," *International Journal of Forecasting*, 9, 387–397.

GLOSTEN, L., R. JAGANNATHAN, AND D. E. RUNKLE (1993): "On the relation between the expected value and the volatility of the nominal excess return on stocks," *Journal of Finance*, 48, 1779–1801.

GONZALEZ-RIVERA, G., T. LEE, AND S. MISHRA (2004): "Forecasting volatility: a reality check based on option pricing, utility function, value-at-risk, and predictive likelihood," *International Journal of Forecasting*, 20, 629–645.

GOURIEROUX, C., AND J. JASIAK (2001): *Financial econometrics*. Princeton, New Jersey.

GRAMMATIKOS, T., AND A. SAUNDERS (1986): "Future price variability: a test of maturity and volume effects," *Journal of Business*, 59, 319–330.

GRANGER, C., AND P. NEWBOLD (1986): *Forecasting economic time series*. Academic Press, San Diego.

GRANGER, C., AND M. PESARAN (2000): "Economic and statistical measures of forecast accuracy," *Journal of Forecasting*, 19, 537–560.

GREGORY, D. (1978): "Multiplicative risk premiums," *Journal of Financial and Quantitative Analysis*, 13, 947–963.

GROSSMAN, S., AND J. STIGLITZ (1980): "On the impossibility of informationally efficient markets," *American Economic Review*, 70, 393–408.

GUO, D. (1996a): "The information content of implied stochastic volatility from currency options," *Canadian Journal of Economics*, 29, 559–561.

——— (1996b): "The predictive power of implied stochastic variance from currency options," *Journal of Futures Markets*, 16, 915–942.

HAMILTON, J., AND G. LIN (1996): "Stock market volatility and the business cycle," *Journal of Applied Econometrics*, 11, 573–593.

HAMMOND, J. (1974): "Simplifying the choice between uncertain prospects where preference is nonlinear," *Management Science*, 20, 1047–1072.

HAUGEN, R. (2003): *Modern investment theory, 5th edn.* Prentice Hall, Upper Saddle River.

HEYNEN, R., AND H. KAT (1994): "Volatility prediction: a comparison of the stochastic volatility, GARCH(1,1) and EGARCH(1,1) models," *Journal of Derivatives*, Summer, 50–65.

HOWELLS, P., AND K. BAIN (2000): *Financial markets and institution.* Pearson Education, England.

HU, M., AND C. TSOUKALAS (1999): "Combining conditional volatility forecasts using neural networks: an application to the EMS exchange rates," *Journal of International Financial Markets*, 9, 407–422.

HULL, J. (2003): *Options, futures, and other derivatives (5th Edition).* Prentice Hall, New Jersey.

JONES, C. (2001): "The dynamics of stochastic volatility: evidence from underlying and options markets," *Working paper, Simon School Business, University of Rochester.*

JORION, P. (1995): "Predicting volatility in the foreign exchange market," *Journal of Finance*, 50, 507–528.

——— (1996): "Risk and turnover in the foreign exchange market," *in The Microstructure of Foreign Exchange Markets, Frankel J.A., Galli G. and Giovannini A. eds.*, National Bureau of Economic Research Conference Report Series, 19–37.

KALLBERG, J. G., AND W. T. ZIEMBA (1983): "Comparison of alternative utility functions in portfolio selection problems," *Management Science*, 29, 1257–1276.

KEENEY, R., AND H. RAIFFA (1976): *Decisions with multiple objectives.* John Wiley and Sons, New York.

Kilian, L., and M. Taylor (2003): "Why is it so difficult to beat the random walk forecast of exchange rates?," *Journal of International Economics*, 60, 85–107.

Kilic, R. (2007): "Conditional volatility and distribution of exchange rates: GARCH and FIGARCH," *Studies in Nonlinear Dynamics and Econometrics*, 11, 1–31.

Knight, J., and S. Satchell (1998): *Forecasting volatility in the financial markets*. Butterworth-Heinemann, Oxford.

Kroner, K. (1996): "Creating and using volatility forecasts," *Derivatives Quarterly*, 3, 39–53.

Lambert, D., and B. McCarl (1985): "Risk modelling using direct solution of nonlinear approximations of the utility function," *American Journal of Agricultural Economics*, 67, 4846–4852.

Latane, H., and R. Rendleman (1976): "Standard deviations of stock price ratios implied in option prices," *Journal of Finance*, 31, 369–381.

LeBaron, B. (2005): *Agent-based computational finance*. the preliminary draft for The Handbook of Computational Economics, edited by K. L. Judd and L. Tesfatsion, vol. II, International Business School, Brandeis University, Waltham, MA.

Lee, K. (1991): "Are the GARCH models best in out-of-sample performance?," *Economics Letters*, 37, 305–308.

Lettau, M., and S. Ludvigson (2001): "Consumption, aggregare wealth, and expected stock returns," *Journal of Finance*, 56, 815–849.

Levy, H., and H. Markowitz (1979): "Approximating expected utility by a function of mean and variance," *American Economic Review*, 69, 308–317.

Li, K. (2002): "Long-memory versus option-implied volatility prediction," *Journal of Derivatives*, 9, 9–25.

Lintner, J. (1965): "The valuation of risky assets and the selection of risky investment in stock portfolios and capital budgets," *Review of Economics and Statistics*, 47, 13–37.

Lopez, J. (2001): "Evaluating the predictive accuracy of volatility models," *Journal of Forecasting*, 20, 87–109.

Lucas, R. (1978): "Asset prices in an exchange economy," *Econometrica*, 46, 1429–1446.

——— (1982): "Interest rates and currency prices in a two-country world," *Journal of Monetary Economics*, 10, 335–360.

Madura, J. (1998): *Financial markets and institutions*. South-Western College Publishing, Ohio.

Mandelbrot, B. (1963): "The variation of certain speculative prices," *Journal of Business*, 36, 394–416.

Martens, M.and Zein, J. (2002): "Predicting financial volatility: high-frequency time-series forecasts vis--vis implied volatility," *Working paper, Erasmus University*.

MATLAB. (2008): *MATLAB 7 User Guides*. The MathWorks, Inc., MA.

McCrae, M., Y. Lin, D. Pavlik, and C. Gulati (2002): "Can cointegration-based forecasting outperform univariate models? An application to Asian exchange rates," *Journal of Forecasting*, 21, 355–380.

McKenzie, M. (1999): "Power transformation and forecasting the magnitude of exchange rate changes," *International Journal of Forecasting*, 15, 49–55.

McQueen, G., and K. Vorkink (2004): "Whence GARCH? a preference-based explanation for conditional volatility," *Review of Financial Studies*, 17, 915–949.

Mehra, R., and E. Prescott (1985): "The equity premium: a puzzle," *Journal of Monetary Economics*, 15, 145–161.

Merton, R. (1973): "The theory of rational option pricing," *Bell Journal of Economics and Management Science*, 4, 141–183.

——— (1980): "On estimating the expected return on the market: an exploratory investigation," *Journal of Financial Economics*, 8, 323361.

Moore, M., and M. Roche (2002): "Less of a puzzle: a new look at the forward forex market," *Journal of International Economics*, 58, 387–411.

——— (2004): "A neo-classical explanation of nominal exchange rate volatility," *Queens University Belfast*.

——— (2006): "Solving exchange rate puzzles with neither sticky prices nor trade costs," *working paper, Queen's University Belfast*.

Morana, C., and A. Beltratti (2004): "Structural change and long-range dependence in volatility of exchange rates: either, neither or both?," *Journal of Empirical Finance*, 11, 629–658.

Neely, C. (1999): "Target Zones and Conditional Volatility: The Role of Realignments," *Journal of Empirical Finance*, 6, 177–192.

——— (2004): "Forecasting foreign exchange volatility: why is implied volatility biased and inefficient? and does it matter?," *Federal Reserve Bank of St. Louis*, Working Papers No. 2002-017.

Nelson, D. (1991): "Conditional heteroscedasticity in asset returns: a new approach," *Econometrica*, 59, 347–370.

Nelson, D., and C. Cao (1992): "Inequality constraints in the univariate GARCH model," *Journal of Business and Economic Statistics*, 10, 229–235.

Noh, J., R. Engle, and A. Kane (1994): "Forecasting volatility and option prices of the SP 500 index," *Journal of Derivatives*, Fall, 17–30.

Pagan, A., and G. Schwert (1990): "Alternative models for conditional stock volatility," *Journal of Econometrics*, 45, 267–290.

Park, B. (2002): "An outlier robust GARCH model and forecasting volatility of exchange rate returns," *Journal of Forecasting*, 21, 381–393.

PARKIN, M. (1970): "Discount house portfolio and debt selection," *The Review of Economic Studies*, 37, 469–497.

PONG, S., M. SHACKLETON, S. TAYLOR, AND X. XU (2002): "Forecasting currency volatility: a comparison of implied volatilities and AR (FI) MA models," *Working paper, Lancaster University.*

PONG, S., M. SHACKLETON, S. TAYLOR, AND X. XU (2004): "Forecasting currency volatility: a comparison of implied volatilities and AR (FI) MA models," *Journal of Banking and Finance*, 28, 2541–2563.

POON, S. (2005): *A practical guide to forecasting financial market volatility.* John Wiley and Sons, England.

POON, S., AND C. GRANGER (2003): "Forecasting volatility in financial markets: a review," *Journal of Economic Literature*, 41, 478–539.

——— (2005): "Practical issues in forecasting volatility," *Financial Analysts Journal,*, 61, 45–56.

PRATT, J. (1964): "Risk aversion in the small and in the large," *Econometrica*, 32, 122–136.

RADALJ, K., AND M. MCALEER (2005): "Speculation and destabilization," *Mathematics and computers in simulation*, 69, 151–161.

RAFFAELLA, G., AND K. IVANA (2005): "Evaluation and combination of conditional quantile forecasts," *Journal of Business and Economic Statistics*, 23, 416–431.

RAVN, M., S. SCHMITT-GROHE, AND M. URIBE (2006): "Deep habits," *Review of Economic Studies*, 73, 1–24.

RODRIGUEZ, J., AND E. RUIZ (2003): "A powerful test for conditional heteroscedasticity for financial time series with highly persistent volatilities," *Universidad Carlos III, Departamento de Estadstica y Econometra*, Statistics and Econometrics Working Papers, no. ws036716.

ROSS, S. (1976): "The arbitrage theory of capital asset pricing," *Journal of Economic Theory*, 13, 341–360.

SARGENT, T. (1987): *Macroeconomic Theory, 2nd edn.* Academic Press, New York.

SARNO, L., AND M. TAYLOR (2002): *The economics of exchange rates.* Cambridge University Press, Cambridge.

SCHWARZ, G. (1978): "Estimating the dimension of a model," *Annals of Statistics*, 6, 461–464.

SCHWERT, G. (1989): "Why does stock market volatility change over time?," *Journalof Finance*, 45, 1115–1153.

SCOTT, E., AND T. A.L. (1989): "Predicting currency return volatility," *Journal of Banking and Finance*, 13, 839–851.

SHARPE, W. (1964): "Capital asset prices: a theory of market equilibrium under conditions of risk," *Journal of Finance*, 19, 425–442.

——— (2007): "Expected utility asset allocation," *Financial Analysts Journal*, 63, 18–30.

Spiro, P. (1990): "The impact of interest rate changes on stock price volatility," *Journal of Portfolio Management*, 16, 63–68.

Still, D. (1993): "Predicting stock-market volatilit," *Business Review, Federal Reserve Bank Philadelphia*, Jan/Feb, 15–27.

Sundaresan, S. (1989): "Intertemporally dependent preferences and the volatility of consumption and wealth," *Review of Financial Studies*, 2, 73–89.

Szakmary, A., E. Ors, J. Kim, and W. Davidson III (2002): "The predictive power of implied volatility: Evidence from 35 futures markets," *Working paper, Southern Illinois University*.

Tambakis, D., and A. Royen (2002): "Conditional predictability of daily exchange rates," *Journal of Forecasting*, 21, 301–315.

Taylor, S. (1986): *Modelling financial time series*. Wiley, Chichester.

——— (1987): "Forecasting of the volatility of currency exchange rates," *International Journal of Forecasting*, 3, 159–170.

——— (2005): *Asset price, dynamics, volatility, and prediction*. Princeton University Press, New Jersey.

Taylor, S., and G. Xu (1997): "The incremental volatility information in one million foreign exchange quotations," *Journal of Empirical Finance*, 4, 317–340.

Theil, H. (1960): *Economic forecasts and policy, 2nd edition*. North-Holland, Amsterdam.

Tims, B., and R. Mahieu (2003): "A range-based multivariate model for exchange rate volatility," *ERIM Report Series Reference*, No. ERS-2003-022-FA.

Tobin, J. (1958): "Liquidity preference as behavior towards risk," *Review of Economic Studies*, 25, 65–86.

Tse, Y. (1991): "Stock returns volatility in the Tokyo Stock Exchange," *Japan and the World Economy*, 3, 285–298.

Tse, Y., and S. Tung (1992): "Forecasting volatlity in the Singapore stock market," *Asia Pacific Journal of Management*, 9, 1–13.

Uribe, M. (2002): "The price-consumption puzzle of currency pegs," *Journal of Monetary Economics*, 80, 38–42.

Vasilellis, G., and N. Meade (1996): "Forecasting volatility for portfolio selection," *Journal of Business Finance and Accounting*, 23, 125–143.

Verdelhan, A. (2006): "A habit based explanation of the exchange rate risk premium," *Computing in Economics and Finance*, 217.

Vilasuso, J. (2002): "Forecasting exchange rate volatility," *Economics Letters*, 76, 5964.

WEI, S., AND J. FRANKEL (1991): "Are option-implied forecasts of exchange rate volatility excessively variable?," *NBER working paper No. 39.*

WEST, K., AND D. CHO (1995): "The predictive ability of several models of exchange rate volatility," *Journal of Econometrics*, 69, 367–391.

WEST, K., H. EDISON, AND D. CHO (1993): "A utility-based comparison of some models of exchange rate volatility," *Journal of International Economics*, 35, 23–45.

WHITE, D. (1966): "Forecasts and decision making," *Journal of Mathematical Analysis and Applications*, 14, 163–173.

WIGGINS, J. (1992): "Estimating the volatility of SP 500 futures prices using the extreme-value method," *Journal of Futures Market*, 12, 256–273.

WRIGHT, J., AND T. BOLLERSLEV (1999): "High frequency data, frequency domain inference and volatility forecasting," *International finance discussion papers, No. 649, Board of Governors of the Federal Reserve System.*

XEKALAKI, E., AND S. DEGIANNAKIS (2005): "Evaluating volatility forecasts in option pricing in the context of a simulated options market," *Computational Statistics and Data Analysis*, 49, 611–629.

XU, X., AND S. TAYLOR (1995): "Conditional volatility and the informational efficiency of the PHLX currency options market," *Journal of Banking and Finance*, 19, 803–821.

ZAKOIAN, J. (1994): "Threshold heteroskedastic models," *Journal of Economic Dynamics and Control*, 18, 931–955.

ZUMBACH, G. (2002): "Volatility processes and volatility forecast with long memory," *Working paper, Olsen Associates.*

ZUMBACH, G., O. PICTET, AND O. MASUTTI (2001): "Genetic programming with syntactic restrictions applied to financial volatility forecasting," *Working Paper, Olsen Associates.*