ESSAYS ON VOLATILITY FORECASTING

Dimos S. Kambouroudis

A Thesis Submitted for the Degree of PhD
at the
University of St. Andrews

2012

Essays on Volatility Forecasting

Dimos S Kambouroudis

This thesis is submitted for the degree of PhD
at the
University of St Andrews

Submitted September 2011

Declarations

1. Candidate's declarations

I, Dimos S Kambouroudis, hereby certify that this thesis, which is approximately 56000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in September 2006 and as a candidate for the degree of PhD in September 2006; the higher study for which this is a record was carried out in the University of St Andrews between 2006 and 2011.

Date **Signature of candidate**

2. Supervisor's declarations

I, hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date **Signature of supervisor**

3. Permission for electronic publication

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the electronic publication of this thesis:

Access to printed copy and electronic publication of thesis through the University of St Andrews.

Date

Signature of candidate **Signature of supervisor**

Essays on Volatility Forecasting

Dimos S Kambouroudis

This thesis is submitted for the degree of PhD
at the
University of St Andrews

Submitted September 2011

Abstract

Stock market volatility has been an important subject in the finance literature for which now an enormous body of research exists. Volatility modelling and forecasting have been in the epicentre of this line of research and although more than a few models have been proposed and key parameters on improving volatility forecasts have been considered, finance research has still to reach a consensus on this topic. This thesis enters the ongoing debate by carrying out empirical investigations by comparing models from the current pool of models as well as exploring and proposing the use of further key parameters in improving the accuracy of volatility modelling and forecasting. The importance of accurately forecasting volatility is paramount for the functioning of the economy and everyone involved in finance activities. For governments, the banking system, institutional and individual investors, researchers and academics, knowledge, understanding and the ability to forecast and proxy volatility accurately is a determining factor for making sound economic decisions. Four are the main contributions of this thesis. First, the findings of a volatility forecasting model comparison reveal that the GARCH genre of models are superior compared to the more 'simple' models and models preferred by practitioners. Second, with the use of backward recursion forecasts we identify the appropriate in-sample length for producing accurate volatility forecasts, a parameter considered for the first time in the finance literature. Third, further model comparisons are conducted within a Value-at-Risk setting between the RiskMetrics model preferred by practitioners, and the more complex GARCH type models, arriving to the conclusion that GARCH type models are dominant. Finally, two further parameters, the Volatility Index (VIX) and Trading Volume, are considered and their contribution is assessed in the modelling and forecasting process of a selection of GARCH type models. We discover that although accuracy is improved upon, GARCH type forecasts are still superior.

Acknowledgements

This thesis is the product of hard work, dedication and many sacrifices and would have not been completed without the help and support of certain individuals. In the next few lines I would like to take the opportunity to thank the people who have stood by me during this journey.

I would first like to thank my supervisor Professor David McMillan for his limitless help, support and patience. David made available his support in a number of ways, not only as a supervisor but as a mentor and friend. A simple thank you would not be enough to express my gratitude to you.

I am most grateful to my parents to whom I owe my deepest gratitude. Thank you for always being there for me, thank you for inspiring me, making me appreciate the value of education and thank you for showing me the way…
"*Now father, you cannot tell me to study any more*…"

Last but not least I would like to thank my wife Melisa and my daughters Ioanna and Marina. Thank you for your patience, for putting up with me during this process, supporting me through the highs and lows, for keeping me sane and for adding value to this achievement. I am grateful to you.

List of contents

1. Introduction

Stock market volatility has been an important subject in the finance literature for which now an enormous body of research exists, especially after the 1987 stock market crash. Volatility modelling and forecasting have been in the epicentre of this line of research and although more than a few models have been proposed and key parameters on improving volatility forecasts have been considered, finance research has still to reach a consensus on this matter. With this thesis we wish to enter the ongoing debate and conduct research by comparing models from the current pool of models as well as explore and propose further key parameters to be considered in improving the accuracy of volatility modelling and forecasting.

Accurately modelling and forecasting volatility is of significant importance for anyone involved in the financial markets. In general, according to Figlewski (2004), the term volatility is associated with risk, and high volatility is thought of as a symptom of market disruption implying that assets and securities are not fairly priced. For example increased volatility will have important implications for investors. Investors may have to alter their investment strategies either by shifting their investment portfolios towards less risky short-term assets, or they could use immunisation strategies for their portfolios. On the other hand policymakers are also affected by increased volatility and would pursue regulatory reforms either by trying to reduce volatility directly or by assisting financial markets and institutions to adapt to increased volatility, Becketti and Sellon (1991).

On the other hand other activities such as risk management, portfolio management and selection, derivative pricing and hedging are examples of activities that would suffer without accurate volatility predictions. More specifically, Engle and Patton (2001) mention: "*A risk manager must know today the likelihood that his portfolio will decline in the future. An option trader will want to know the volatility that can be expected over the future of the life of the contract. To hedge this contract he will also want to know how volatile is this forecast volatility. A portfolio manager may want to sell a stock or a portfolio before it becomes more volatile. A market maker may want to set the bid ask spread wider when the future is believed to be more volatile*" p. 2.

As can be seen, the importance of accurately forecasting volatility is paramount for the functioning of the economy and everyone involved in finance activities. In periods of instability, volatility forecasting becomes even more important since governments, the banking system, institutional and individual investors are trying to cope with increased risk, increased volatility and lack of resources. Knowledge, understanding and the ability to forecast and proxy volatility accurately could be a determining factor for survival not only during turbulent times but also during periods of economic growth, giving an advantage to whoever can successfully manage future volatility. This is also the motivation of this thesis.

With this thesis we wish to add knowledge to the literature on volatility forecasting by means of empirical investigation and propose the use of a number of key parameters that should be considered in the modelling process of volatility forecasting. There are four main contributions of this thesis. First, after performing a volatility exercise it is established that the GARCH genre of models are superior to the more 'simple' models

and models used by practitioners. Second, for the first time in the finance literature the question of identifying appropriate in-sample lengths for out-of-sample forecasts is raised. The answer to this question is proven to support the views raised by practitioners, that large in-sample periods are not necessary for producing accurate volatility forecasts. Third, within a Value-at-Risk volatility forecasting setting the RiskMetrics model, which is preferred by practitioners for its simplicity, does a poorer job compared to the more complex GARCH type models. Fourth, the contribution of the Volatility Index (VIX) and Trading Volume on the forecast ability of a selection of GARCH type models cannot be ignored since a better level of accuracy is achieved; however GARCH type forecasts are dominant.

Apart from the main contributions mentioned above other issues are also examined. The sample selection, with the exception of the last empirical chapter where due to data availability only a small number of countries are considered, consists of large number of countries with a good mix of both developed and emerging economies in order to identify any trends, patterns or other regularities. Furthermore, the nature of the topic allowed for comparisons of methods between those used mainly in academia and methods preferred by finance practitioners.

The structure of the thesis comprises of a literature review chapter, four empirical chapters and a conclusion chapter. The next chapter (**Chapter 2**) is the literature review chapter. Definitions of stock market volatility are given and its importance for the finance literature highlighted. The main focus of this chapter is on the exploration of the different models used for forecasting volatility, looking first at the 'simple' models and then after addressing the empirical regularities found in datasets such as

volatility clustering, the leverage effect and stationarity then the more advanced GARCH type models are introduced. The models introduced in this chapter are used in the rest of the thesis. Then follows a section on the state of the literature setting the scene for the research questions we address in this thesis.

In **Chapter 3** a straightforward comparison exercise of volatility models is performed. This type of exercise has been a popular theme within the finance literature with often conflicting results. Stock market volatility has been the subject of numerous studies in the finance literature -see literature review chapter for more details, particularly after the stock market crash of 1987. Similarly, modelling and forecasting volatility became a popular area of research within finance, for academics and practitioners alike. Taking part in this debate a comparison between representative models from the two popular model categories, the 'simple' models; namely the Exponential Smoothing and the Moving Averages and the more 'advanced' GARCH type models capturing the features of volatility clustering, the leverage effect and volatility persistence, which are found to exist in the data.

In an attempt to identify any possible global trends the sample is selected from a wide geographical perspective (Europe, Asia, America and Australia) including developed and emerging markets, since the majority of the empirical work has been carried out mainly on developed markets. For all the countries of the sample daily closing prices of the countries representative index spanning over two decades are obtained. Four measures of comparison are used in this exercise and a further dimension is explored based on the classification of the sample markets in order to identify the existence or not of any differences between emerging and developed economies. The results show

that the more advanced GARCH type models do a better job overall than the 'simple' models. More specifically in the order of the asymmetric models first followed by the long memory models and finally in third place the 'simpler' time series models. When the country classification is taken into account a clearer picture emerges in the ranking of the results for the developed economies than for the developing economies, however for both the developed and emerging economies there is no contest in identifying the worst performing model as the exponential smoothing model.

Chapter 4 takes a look at a key parameter ignored so far by academic research within the volatility forecasting literature, and that is specifying the ideal size of the in-sample period required for producing accurate forecasts. The question of 'how much previous data do we need in order to produce accurate forecasts?' This question introduces the notion of recursive forecasts for the first time within volatility forecasting where the debate is between practitioner/investors and researchers/academics who share different views regarding this question. Respectively, a small in-sample period (small number of observations) is preferred to a large in-sample period (large number of observations) when forecasting volatility due to cost and storage restrictions.

The same dataset from Chapter 3 is used and a good selection of the better performing models from the same chapter are selected, more specifically from the GARCH genre; GARCH(1,1), EGARCH, TGARCH and CGARCH and the representative 'simple' model Moving Average. The main objective of this exercise is to determine the optimal number of in-sample observations to produce the most accurate forecasts. For each model a 'rolling window' of 60 observations is used which rolled back from the

fixed end date to the start of the variable in-sample period producing a forecast for every window of 60 observations. For each forecast (window) the forecast comparison measure of Mincer-Zarnowitz (MZ) performed, regressing the true volatility value on the produced forecast value obtaining the coefficient of determination for comparison purposes.

The results show a degree of homogeneity. For most countries of the sample and for the majority of the models large in-sample periods are not necessary for producing the most accurate forecasts supporting the practitioners/investors view; however the models that produce the most accurate forecasts require larger in-sample durations. Furthermore, when taking into account the country classification smaller in-sample durations are required for producing accurate forecasts in emerging markets but more accurate forecasts produced for countries in developed economies.

The superiority of the GARCH genre of models has been highlighted by the finance literature (see Chapters 2 & 3), however, the aim of **Chapter 5** is to seek an answer to the question whether the in-sample superiority of the GARCH model carries over to out-of-sample forecasting, or whether forecasts from the RiskMetrics model known for its simplicity of application and is preferred by the finance professionals can provide adequate forecasts of volatility in a Value-at-Risk setting.

In the academic finance literature the problems associated with the RiskMetrics model have been reported, more specifically with respect to the undefined unconditional variance and the model's inability to produce long-horizon forecasts. On the other hand the GARCH genre of models has found support by the academic finance

literature, which not only does not suffer from the same problems the RiskMetrics approach does but also is better able to capture the inherent time-dependency within volatility.

Using a large selection of thirty-one international stock markets including those of the G7, thirteen further European markets and eleven further Asian markets RiskMetrics forecasts were compared to those of the GARCH type models within a VaR framework. The following conclusions are reached. When forecasting the 1% VaR the RiskMetrics model does a poor job and is typically the worst performing model, on the other hand the GARCH type models and more specifically the APARCH model is preferred. However when forecasting at the 5% VaR then the RiskMetrics model performs adequately. In short, the RiskMetrics model only performs well in forecasting the volatility of small emerging markets and for broader VaR measures. This chapter and aspects from chapter 3 were published[1] in the International Review of Financial Analysis, a copy can be found in the appendix of the thesis.

In the final empirical chapter (**Chapter 6**) we assess the effect of the Volatility Index (VIX) and Trading Volume on volatility forecasting. Both VIX and Volume have appealing and useful properties, which the finance literature has recognised, resulting in both these factors to be considered in forecasting exercises mainly individually, and with only a very small number of recent studies assessing the impact of both VIX and Volume together within the context of volatility forecasting.

[1] Reference: McMillan, D. G., and Kambouroudis, D., (2009), "Are RiskMetrics forecasts good enough? Evidence from 31 stock markets", *International Review of Financial Analysis*, Vol. 18, pp. 117-124.

VIX has proven to be a useful instrument for forecasting volatility, since it is a forward looking measure and is defined as a benchmark of expected short-term market volatility upon which futures and options contracts on volatility can be written. On the other hand Trading Volume is caused by information flow which is positively correlated to price changes suggesting that a relationship between Trading Volume and volatility also exists.

Following on from the previous empirical chapters VIX and Volume data are used within a GARCH type model framework and the testing procedure of Mincer-Zarnowitz (MZ) is used followed by forecast encompassing tests in order to establish whether there is added value in incorporating the two parameters within the forecasting process. Three main markets selected are the UK, France and the USA mainly due to data availability. The results suggest that both VIX and Volume improve on the informational content of the GARCH type models, VIX does a better job in this process than Volume, but better results are reported when VIX and Volume are used together. In answering the question whether VIX produces better forecasts than the GARCH genre of models, the answer is no but the informational content of VIX cannot be ignored.

Finally, **Chapter 7** summarises the main findings: the superiority of the GARCH genre of models in volatility forecasting exercises over the 'simpler' time series models, models preferred by finance practitioners and VIX; and that the in-sample duration is an important determinant for out-of-sample forecasts. This chapter also provides concluding remarks as well as propositions for future research on the issues addressed in this thesis.

Note: The use of the first person plural ‘we’ instead of the first person singular ‘I’ is used throughout the thesis. Any work published from this thesis will be co authored jointly with my supervisor Professor David McMillan.

2. Literature review

2.1 Introduction

Stock market volatility has been the subject of many studies over the past few decades. The main impetus for this interest began after the 1987 stock market crash where, for example, the Standard & Poor's (S&P) composite portfolio dropped from 282.70 to 224.84 (20.4 %) and the Dow-Jones Average fell by 508 points in one day.[2] The term stock market volatility refers to the characteristic of the stock market to rise or fall sharply in price within a short-term period (from day to day or week to week). A complete definition of volatility in the economic sense is given by Andersen et al. (2005) in a more recent working paper: "*Volatility within economics is used slightly more formally to describe without a specific implied metric, the variability of the random variable (unforeseen) component of a time series. More precisely, in financial econometrics, volatility is often defined as the (instantaneous) standard deviation (or σ "sigma") of the random Wiener-driven component in a continuous-time diffusion model. Expressions such as "implied volatility" from option prices rely on this terminology*" (p.1). This phenomenon is not new since throughout the post-war period, stock markets, commodity markets, bond markets and foreign exchange markets have recorded sharp movements.

In a review essay by Cochrane (1991) several issues are being addressed: "*what, ultimately, is behind day-to-day movements in prices? Can we trace the source of movements back in a logical manner to fundamental shocks affecting the economy…?*

[2] October 19, 1987 was the largest percentage change in market value in over 29,000 days. Stock volatility jumped dramatically during and after the crash (Schwert, 1990).

Are price movements due to changes in opinion or psychology, that is, changes in confidence, speculative enthusiasm…?" (p. 463). Furthermore economists and other academics were concerned about the efficient market hypothesis and volatility; do volatility tests reject the efficiency itself? As Shiller (1989) and Cochrane (1991) mention, volatility tests do not prove that markets are inefficient. "*Volatility tests are in fact only tests of specific discount-rate models, and they are equivalent to conventional return-forecasting tests… Thus, the bottom line of volatility tests is not 'markets are inefficient' since 'prices are too volatile', but simply 'current discount-rate models leave a residual' since (discounted) returns are forecastable"*, Cochrane (1991), p. 464.

The determinants of financial market volatility, according to Shiller (1988) are difficult to define, simply because economists and other researchers do not have a proven theory of financial fluctuations. Theories that exist are often unconvincing. One explanation of financial market volatility, given by Shiller (1988), is market psychology. Investors appear sometimes to react to each other instead to some fundamental event, and this process can set into motion large market swings. He proved with his survey that market psychology was a key factor behind the stock market crash of 1987, suggesting that on the day of the crash investors were not responding to any specific news item but to news of the crash itself. Mishkin (1988) agreed with Shiller that stock market volatility is difficult to explain, and although he did not fully agree with his survey evidence, he too believed that factors other than underlying economic fundamentals may have played a role in the stock market crash of 1987.

The notion of speculation has been mentioned and was related to the 'bad' effects of volatility.[3] There is a debate about speculators and their impact on volatility, suggesting that increased volatility is undesirable and reductions in volatility are desirable. This is misleading as it fails to recognise the link between information and volatility, Antoniou and Holmes (1995). Within the Efficient Market Hypothesis (EMH) literature there is a positive relationship, with a rapid reaction between the arrival of information and price fluctuations. Consequently, if the flow of information increases, in an efficient market, price movements will be more frequent (more volatile), Antoniou et al. (1997).

Generally, increased volatility has been viewed as an undesirable consequence of destabilising market forces such as speculative activity, noise trading or feedback trading. Increased volatility could come as a result of an innovation, by reflecting the actual variability of information regarding fundamental values. So increased volatility may not necessarily be undesirable, Bollerslev et al. (1992).

Ross (1989) using a simple model under the condition for no arbitrage, proved that the variance of price change will be equal to the rate (or variance) of information flow. *"In an arbitrage-free economy, the volatility of prices is directly related to the rate of flow of information to the market. In a simple model the two were found to be identical. This result links volatility tests to efficient market hypothesis which specify the information set the market uses for pricing"* (p. 17). By this we can conclude that the volatility of the asset price and in consequence the volatility of the market as a

[3] It is argued that speculators can have a destabilising impact on prices De Long, J.B., Shleifer, A., Summers, L.H. & Waldmann, R.J., (1990).

whole, will increase as the rate of information increases. In the opposite case arbitrage opportunities will exist.

Mishkin (1988) also addressed the role of monetary policy in the face of financial market volatility. Monetary policymakers have two alternatives when dealing with volatility. They can attempt to reduce the volatility by intervening in markets, or they can stay out of the markets but stand ready to function as lender of last resort in the event of a financial crisis. He indicated a preference for the latter.

A stock market fall could be harmful for the economy. It has been observed, according to Becketti and Sellon (1991), that stock volatility has an effect on the economy through consumer spending, business investment spending and also could disrupt the smooth functioning of the financial system by leading to structural regulatory changes.

First, stock price volatility hinders the performance of the economy via consumer spending. Immediately after the October 1987 drop in stock prices, economic forecasts predicted sharply weaker economic growth. It was believed that the fall in stock prices would reduce consumer spending, because of the weakening of consumer confidence and wealth. Second, investors may perceive a rise in the stock market volatility as an increase in the risk of equity investments. Thus, investors could shift their funds to less risky assets i.e. bonds, although long-term investments contain an element of risk too. This reaction would tend to raise the cost of equity for firms issuing stock and to misallocation of resources Antoniou et al. (1997). Small and new firms could suffer as a consequence of the effect, since investors will move toward the

purchase of stock in large and well-known firms. Finally, extreme stock price movements could also have an effect on the financial mechanism and lead to structural changes. Systems working under normal price volatility may be unable to cope with extreme price changes. The system itself may contribute to volatility if investors are unable to complete stock transactions. Changes in market rules or regulations may be necessary to increase the resilience of the market in the face of greater volatility, Becketti and Sellon (1991).

Changes in volatility have important implications for investors and policymakers. Investors may have to alter their investment strategies. They would have two alternatives in order to cope with increased volatility. They could either shift their investment portfolios towards less risky short-term assets, or they could use immunisation, for example hedging or other strategies, for their portfolios. For instance, investors after the October 1987 crash tried to adjust to volatility by restructuring their portfolios. This explains the sharp drop in stock purchases after the crash. Individual investors reduced their direct purchases of stocks and also shifted away from stock mutual funds. As a consequence, retail stock brokerages and mutual funds have experienced reduced profitability and have scaled back operations and employment.

On the other hand policymakers may pursue regulatory reforms by either trying to reduce volatility directly or by assisting financial markets and institutions to adapt to increased volatility. In practice policymakers have focused on the latter, improving the ability of financial markets and institutions to weather increased volatility. For financial institutions directly exposed to increased volatility, such as depository

institutions and market makers, policymakers have encouraged greater capitalisation. Increased capital allows these institutions to weather greater financial volatility without incurring the liquidity and solvency problems that might disrupt the functioning of financial markets, Becketti and Sellon (1991).

The topic of volatility is of significant importance to anyone involved in the financial markets. In general volatility has been associated with risk, and high volatility is thought of as a symptom of market disruption, with securities unfairly priced and the malfunctioning of the market as whole. Especially within the derivative security market volatility and volatility forecasting is vital as managing the exposure of investment portfolios is crucial, Figlewski (2004). More recently the literature has focused on the ability to forecast volatility of asset returns. There are many reasons why forecasting volatility is important according to Walsh, Yu-Gen Tsou (1998), for example, option pricing has traditionally suffered without accurate volatility forecasts. Controlling for estimation error in portfolios constructed to minimise *ex ante* risk, with accurate forecasts we have the ability to take advantage of the correlation structure between assets. Finally when building and understanding asset pricing models we must take into account the nature of volatility and its ability to be forecasted, since risk preferences will be based on market assessment of volatility.

It is apparent that volatility is important, since it directly and indirectly affects the financial system and the economy as a whole. The main aim of this chapter will be to look into the aspects of volatility forecasting since in view of all the reasons explained above researchers, policymakers and investors will have an advantage when accounting for risk, determining economic strategies and making profits. A further

insight into volatility forecasting will be given but first it is necessary to explore the different models used when estimating volatility.

2.2 Modelling Volatility

The complex issue of measuring and quantifying volatility is still one of the main challenges academics and practitioners are dealing with. Over the years several models and methods have been proposed but still we are far from any generally accepted formula. In the next section some of the earlier applications and models of volatility measurement are described.

2.2.1 Early applications for volatility measurement

In order to quantify and model volatility we first must define volatility. In his work, Figlewski (2004), does a good job in setting the groundwork for understanding the concept of volatility using finance basics. Starting from the 'efficient markets' or 'random walk' model, asset price movements can be described by an equation like:

$$r_t = \frac{S_t - S_{t-1}}{S_t} = \mu_t + \varepsilon_t \qquad (2.1)$$

where; $E[e_t] = 0$ and $Var[e_t] = \sigma_t^2$

He mentions, *"the return at time t, r_t, is the percentage change in the asset price S, over the period from t-1 to t. This is equal to μ_t, a non-random mean return for period t, plus a zero mean random disturbance ε_t, that is independent of all past and*

future ε_t *'s. It is the lack of serial correlation in the random* ε_t *'s that is the defining characteristic of efficient market pricing: past price movements give no information about the sign of the random component of return in period t,* Figlewski (2004), p. 3.

Modern option pricing theory began in 1973 with Black and Scholes (1973), where volatility plays a central role in determining the fair value for an option or a derivative instrument with option features. The input parameters required in order to produce a Black–Scholes option price are: the current stock price, the option strike price, the risk free interest rate, the option's remaining time to maturity and the future volatility of the underlying asset. All parameters except the latter one can be obtained easily by the market Figlewski (2004).

$$dS/S = \mu\, dt + \sigma\, dz \qquad (2.2)$$

Options and futures benefit from price fluctuations of the underlying asset as well as of securing a portfolio against price losses or hedging a planned purchase against a possible price increase Maris et al. (2004). In deriving the option pricing formula, Black and Scholes needed to model stock price movements over very short time intervals in order to adjust their trading strategy after continually rebalancing a portfolio consisting of an option and its underlying stock. The formula they adopted (equation 2.2) is a logical extension of the random walk model over time. This is a limiting random walk process as the time interval goes to zero, keeping the mean and the variance of returns per year constant. The result is a lognormal diffusion model where the *dS* is the asset price change over time (infinitesimal time) interval dt, μ is the mean return at an annual rate, *dz* is a time independent random disturbance term

with mean of zero and variance of one at *dt*, and σ is the volatility, i.e. the standard deviation of the annual return, Figlewski (2004).

In order for empirical research to analyse and draw conclusions, an array of models for measuring volatility were developed. Volatility modelling and forecasting has been the subject of vast empirical and theoretical research over the past decades as volatility has become one of the most important concepts in the finance literature. Often volatility is measured by the standard deviation or variance of returns as a simple risk measure. Other models such as Value at Risk (VaR) modelling, for measuring market risk, and the previously mentioned Black and Scholes model, for pricing options, require the estimation of volatility. We consider the returns process given by:

$$r_t = m_t + \varepsilon_t \tag{2.3}$$

where m_t is the conditional mean process (which could include autoregressive (AR) and moving average (MA) terms), where the error term can be decomposed as $\varepsilon_t = \sigma_t z_t$ with z_t an idiosyncratic zero-mean and constant variance noise term, and σ_t is the volatility process to be estimated and forecast, with forecast values denoted h_t^2. The sample data is split between the in-sample period, $t=1,\ldots,T$, and the out-of-sample period $t=T,\ldots,\tau$. In order to generate a historical 'actual volatility' series on the basis of which volatility forecasts may be generated using the statistical models described below, the methodology by Pagan and Schwert (1990) is followed in representing past volatility by the squared residuals from a conditional mean model, for returns estimated over the in-sample period.

2.2.2 Simple Models

The term 'simple' for the models described below refers to the traditional and widely used in the past techniques not only in finance but other disciplines too. The list of models belonging in this category is extensive and only few models will be described here.

Random Walk

If volatility fluctuates randomly the optimal forecast of next period's volatility is simply current actual volatility:

$$h_{t+1}^2 = \sigma_t^2 \tag{2.4}$$

This random walk model thus suggests that the optimal forecast of volatility is for no change since the last observed value.

Historical Average

Extrapolation of the historical mean of the volatility process is perhaps the most obvious means of forecasting future volatility. Furthermore, if the distribution of volatility has a constant mean all variation in estimated volatility could be attributed to measurement error and the historical mean computed below gives an optimal forecast for all future periods:

$$h_{t+1}^2 = \frac{1}{\tau - T} \sum_{t=1}^{\tau} \sigma_t^2 \tag{2.5}$$

Simple Moving Averages

Under the moving average method volatility is forecast by an unweighted average of previously observed volatilities over a particular historical time interval of fixed length:

$$h_{t+1}^2 = \frac{1}{p}\sum_{j=1}^{P}\sigma_j^2 \qquad (2.6)$$

where P is the moving average period or 'rolling window'. The choice of this interval is arbitrary.

Exponential Smoothing

Under exponential smoothing the one-step ahead volatility forecast is a weighted function of the immediately preceding volatility forecast and actual volatility:

$$h_{t+1}^2 = \phi h_t^2 + (1-\phi)\sigma_t^2 \qquad (2.7)$$

where ϕ is a smoothing parameter constrained to lie between zero and one, such that for $\phi=0$, the exponential smoothing model reduces to a random walk model, while for $\phi=1$ weight is given only to the prior period forecast. The value of ϕ is determined empirically by that value which minimises the in-sample sum of squared prediction errors. Empirical studies have also confirmed the usefulness of the exponential smoothing model, Boudoukh et al. (1997).

In 1989 JP Morgan developed the RiskMetrics approach to volatility using the simple exponential smoothing model described above in order to quantify and assess the risk exposure of the firm. This approach received wide acceptance in the finance world as well as in academia when in 1992 JP Morgan launched the RiskMetrics methodology. Since its inception several versions of the RiskMetrics Technical Document have been published, in addition the establishment of the Value at Risk framework also triggered the popularity of the RiskMetrics approach. Empirical studies have also confirmed the usefulness.

Exponentially Weighted Moving Average (EWMA)

The exponentially weighted moving average model is similar to the exponential smoothing model discussed above, but past observed volatility is replaced with a moving average forecast, as in the simple moving average model:

$$h_{t+1}^2 = \phi h_t^2 + (1-\phi)\frac{1}{p}\sum_{j=1}^{P}\sigma_j^2 \qquad (2.8)$$

Exponentially weighted moving average models (EWMA) are an extension of the historical volatility measure allowing more recent observations having a stronger influence on volatility forecasting than older observations. When applying the EWMA modelling the latest observation carries the largest weight and weights associated with previous observations decline exponentially over time. In contrast to the simple historical volatility model, volatility is affected more by recent events which carry more weight than events further in the past and the effect of a single given observation declines at exponential rate.

Smooth Transition Exponential Smoothing method[4]

These models allow parameters to change over time in order to adapt to changes in the characteristics of the time series. Taylor (2004) proposes the use of logistic function of a user –specified variable adaptive smoothing parameter. Within the volatility forecasting modelling framework this would be formulated as:

$$h_{t+1}^2 = a_t + \varepsilon_t^2 + (1 - a_t)h_t^2 \qquad (2.9)$$

where; $a = 1/(1 + \exp(\beta + \gamma V_t))$.

The smoothing parameter varies between zero and one, and adapts to changes in the transition variable V_t and where ε_t and $|\varepsilon_t|$ are used as transition variables in the similar way the sign and size of past shocks have been used as transition variables in non-linear GARCH models, Taylor (2001). In the words of Granger and Poon (2003) the smooth transition model is a more flexible version of the exponential smoothing model where the weight depends on the size and sign of the previous return. This approach is analogous to the GARCH type models for allowing the dynamics of the conditional variance to be influenced by the 'leverage effect' and the 'volatility persistence', characteristics found in stock market data and discussed in more depth in section 2.3.

Other volatility models

Due to the popularity of the topic volatility modelling forecasting several other models have been proposed, this list is also extensive and for this reason only a small representative selection of alternative modes are mentioned below. It was previously

[4] See Taylor (2004)

mentioned that a simple measure of volatility is the standard deviation of returns of an index. Becketti and Sellon (1991) measure volatility by the annual standard deviation of monthly returns in the S&P 500 Composite Stock Price Index. This is a measure of dispersion of monthly returns about the average return for each year. Another method to measure normal volatility mentioned from the same source (Becketti and Sellon, 1991) is by using the interquartile range, the distance between 25th and 75th percentile of the monthly returns within a year.[5]

In a more recent review paper Poon and Granger (2003) attempt to make a distinction between the standard deviation, volatility and risk. Standard deviation, σ, or variance, σ^2, is computed from a set of observations as:

$$\hat{\sigma}^2 = \frac{1}{N-1}\sum_{t=1}^{N}(R_t - \bar{R})^2 \qquad (2.10)$$

where $\bar{R}$ is the mean return. The argument here is that the standard deviation is only the correct dispersion measure for the normal dispersion measure for the normal distribution. The link between volatility and risk is a questionable one. Risk is usually associated with small or negative returns, whereas most measures of dispersion make no distinction. The two examples mentioned by Poon and Granger (2003), are: first is the Sharpe ratio, defined as the return in excess of risk free rate divided by the standard deviation which is frequently used as an investment performance measure occasionally penalizes occasional high returns. Second the ‘semi-variance’, a concept developed by Harry Markowitz (1991), where only the squared returns below the mean are used, but this method is not easy to apply and it is not widely used.

[5] Normal volatility refers to the ordinary variability of stock returns, the ordinary ups and downs in returns.

Grabel (1995) uses two general models the Keynesian Volatility Index and the Neo-Classical Volatility Index (type 1, 2). These indices derive from the theory behind each approach, which are respectively: "*…that assets yield some normal return over time based on their underlying fundamental value. The magnitude of the deviation from the asset's fundamentals-based return constitutes asset volatility"* and *"volatility in the Keynesian case is simply given by the magnitude of asset return fluctuations"* (p. 906, Grabel, 1995).

2.3 Stochastic Volatility (SV)

Stochastic Volatility (SV) models have their roots in mathematical finance and financial econometrics. Interest in this class of models dates at least to the work of Clark (1973) where as suggested modelling asset returns as a function of a random process of information arrival. This so-called time deformation approach yielded a time-varying volatility model of asset returns (Chysels, Harvey and Renault, 1996). Tauchen and Pitts (1991) noted that if the information flows are autocorrelated, then a stochastic volatility model with time varying and autocorrelated conditional variance might be appropriate for price-change series, linking information arrival to asset returns. A different view was expressed by Hull and White (1987) and Melino and Turnbull (1990) where stochastic volatility models could also arise as discrete approximations to various diffusion processes of interest. For example as mentioned in Chysels et al. (1996), they were not directly concerned linking asset returns to information arrival but in pricing European Options assuming continuous time SV

models for the underlying asset. Taylor (1986) on the other hand formulated a discrete time series SV model as an alternative to ARCH models instead of using a likelihood-based approach but the Method of Moments (MM) in order to avoid integration problems associated with evaluating the likelihood directly. As mentioned in Granger and Poon (2005) the volatility noise term makes the SV model a lot more flexible, but as a result the SV model has no closed form hence making Maximum Likelihood unsuitable. In addition to the MM approach other estimation approaches were proposed such an example would be the Quasi-maximum likelihood (QML) estimator of Harvey et al (1994) nevertheless if volatility proxies are non-Gaussian this method is also inefficient. Other alternatives are variations of the Generalised Method of Moments (GMM) approach through simulations, analytical solutions, and the likelihood approach through numerical integration[6]. The Stochastic Volatility model is defined as:

$$R_t = \mu_i + \varepsilon_t \tag{2.11}$$

where $\varepsilon_t = \zeta_t \exp(0.5h_t)$ and $h_t = \omega + \beta h_{t-1} + \upsilon_t$.

Note: υ_t may or may not be independent of ζ_t

[6] A number of studies can be mentioned here as cited in Granger and Poon (2005).

2.4 Implied Volatility (IV)

As previously mentioned in section 2.2.1, the Black-Scholes option pricing formula states that the option price is a function of the price of the underlying asset, the strike price, the risk free interest rate, the time to option maturity and the volatility of the underlying asset. Given that the above parameters are observable, once the market has produced a price for the option, volatility could be derived using backward induction, and then use the volatility measure (value) that the market used as input. This measure of volatility is called option implied volatility. Option implied volatility is often interpreted as a market's expectation of volatility over the option's maturity. Because each asset can have only one volatility measure, difficulties arise when options with similar maturities but different strikes produce different implied volatility estimates for the same asset, Granger and Poon (2005). Some examples of earlier studies where the basic Black-Scholes option pricing model was used are Latane and Rendleman (1976), Chiras and Manaster (1978) and Beckers (1981). More studies also examined implied volatility as a source of information, for examples studies by Day and Lewis (1990) who conclude that time-series models of conditional volatility outperforms implied volatility, and Lamoureux and Lastrapes (1993) who find that information contained in historical volatility is superior to that contained in implied volatility. Furthermore Canina and Figlewski (1993) also find that implied volatility has no correlation with future volatility and hence *"to measure the "market's" volatility estimate (we) must not just take the implied volatility"* (pp. 678).

On the other hand a good number of studies can be found in the finance literature where the superiority of implied volatility is supported, for example, work by Jorion (1995), Fleming (1998), Christensen and Prabhala (1998), and Christensen and

Hansen (2002). More recently, Christensen, Hansen and Prabhala (2001), Blair, Poon and Taylor (2001), Ederinton and Guan (2002), Pong et al. (2004) and Jiang and Tian (2005) take longer data sets, high frequency data and account for structural changes concluding that implied volatility is a more efficient forecast for future volatility than historical volatility.

In the most recent literature review paper by Granger and Poon (2005), it is concluded that the predictive ability of implied volatility cannot be ignored nor underestimated compared to other volatility models. More specifically "*implied volatility appears to have superior forecasting capability, outperforming many historical price volatility models and matching the performance of forecasts generated from time series models that use a large amount of high frequency data*" (pp. 489-490). The importance of implied volatility is highlighted in Chapter 6.

2.5 Empirical regularities of asset returns[7]

Following the seminal work of Mandelbrot (1963) and Fama (1965), many researchers have reported that the empirical distribution of stock returns is significantly non-normal. In particular, the kurtosis of the stock returns time series appears to be larger than the kurtosis of the normal distribution (the time series of stock returns are leptokurtic), the distribution of stock returns can be skewed either to the right (positive skewness) or to the left (negative skewness), and the variance of the stock returns may not be constant over time and indeed volatility exhibits clustering.

[7]This section is mainly based on the work by Bollerslev et al. (1994).

Researchers regarded this as the persistency of the stock market volatility and the financial analyst called this uncertainty or risk. Volatility measured by the variance and covariance and was accepted for decades Chong et al. (1999). In order to select an appropriate volatility model, we must have a good idea of what empirical regularities the model should capture. Some of the important regularities for asset returns are presented below (Bollerslev et al. 1994).

Thick tails

Asset returns tend to be leptokurtic. According to the literature on the modelling of stock returns, stock returns have thick-tailed distributions and are modelled as independent and identically distributed (iid). This empirical regularity has been documented by Mandelbrot (1963), Fama (1963, 1965), Clark (1973), and Blattberg and Gonedes (1974).

Volatility Clustering

Volatility clustering where ‘large changes tend to be followed by large changes of either sign, and small changes tend to be followed by small changes’ as Mandelbrot (1963) wrote, is a visible phenomenon when asset returns are plotted through time. Volatility clustering and thick tailed returns are closely related.

A Sample Financial Asset Returns Time Series

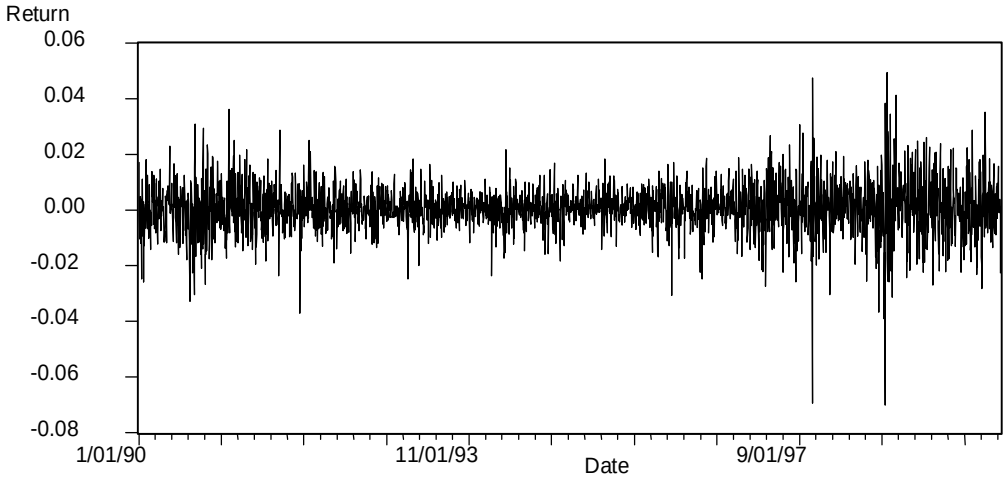


Figure 2.1 Source: Brooks (2008)

Leverage effects

The ‘leverage effect’ noted by Black (1976), refers to the negative correlation between stock prices and changes in stock volatility. Fixed costs such as financial and operating leverage provide a partial explanation for this phenomenon. A firm with debt and outstanding equity usually becomes more highly leveraged when the value of the firm falls. This raises equity returns volatility if the returns on the firm as a whole are constant. However, as argued by Black (1976), the response of stock volatility to the direction of returns is too large to be explained by leverage alone; Christie (1982) and Schwert (1989).

Non-trading periods

Information that accumulates when financial markets are closed is reflected when the markets reopen. For example, information accumulating at a constant rate over time, then the variance of returns over a period from the Friday close to the Monday close should be three times the variance from Monday close to Tuesday close. According to

Fama (1965) and French and Roll (1986), information accumulates at a slower rate when markets are closed than when they are open. Variances are higher after weekends and holidays but not as high as if the rate of information was constant. French and Roll (1986) for example have found that volatility is 70 times higher per hour on average when the market is open than when it is closed, and similar results found by Baillie and Bollerslev (1989) on foreign exchange rates.

Forecastable events

Forecastable releases of important information are associated with high ex ante volatility. For example, Cornell (1978), Patell and Wolfson (1979, 1981) show that individual firms' stock returns volatility is high around earnings announcements, and Harvey and Huang (1991, 1992) find that fixed income and foreign exchange volatility is higher during periods of heavy trading by central banks or when macroeconomic announcements are made. There are also intraday predictable changes in volatility. Several papers have found that volatility is much higher at the opening and closing of a trading day for stocks and foreign exchange; Harris (1986), Gerity and Mulherin (1992) and Baillie and Bollerslev (1991). One explanation for high volatility at the open is the accumulated information, as mentioned above, but the high volatility at closing is more difficult to explain.

Volatility and serial correlation

Both LeBaron (1992) and Kim (1989) find a strong inverse relationship between volatility and serial correlation for the U.S. stock indices and for the foreign exchange respectively. The above finding appears to be remarkably robust to the choice of sample period, market index, measurement interval and volatility measure.

Co-movements in volatilities

Black (1976), mentions that in general when volatilities change, they all tend to change in the same direction. Several studies[8] support this argument of the existence of common factors explaining volatility movements in stock and exchange rates. Engle et al. (1990) show that US bond volatility changes are closely linked across maturities. The commonality of volatility changes holds not only across assets within a market, but also across different markets. For example, it was found by Schwert (1989) that U.S. stock and bond volatilities move together and Engle and Susmel (1993) and Hamao et al. (1990) have discovered links between volatility changes across international stock markets.[9]

As Bollerslev et al. (1994) mention, the fact that volatilities move together should be encouraging to model builders, since it indicates that a few common factors may explain much of the temporal variation in the conditional variances and covariances of asset returns, which is the basis of the ARCH modelling.

Macroeconomic variables and volatility

Stock values are considered to be related and closely tied to the health of the economy, so it is natural to expect that measures of economic uncertainty such as conditional variances of industrial production, interest rates, and money growth, should help explain changes in stock market volatility Bollerslev et al. (1994). Schwert (1989), finds that although stock volatility rises sharply during recessions and financial crises and drops during expansions, the relation between macroeconomic

[8] Diebold and Nerlove (1989) and Harvey et al. (1992).
[9] The importance of international linkages have been further explored by King et al. (1994), Engle et al. (1990) and Lin et al. (1994).

uncertainty and stock volatility is surprisingly weak. On the other hand Glosten et al. (1993), uncover a strong positive relationship between stock return volatility and interest rates.

Long memory

Stock market returns contain little serial correlation, Fama (1970) and Taylor (1986), which complies with the Efficient Market Hypothesis. However, according to Ding et al. (1993), this empirical fact does not suggest that returns are independently identically distributed. More specifically they find that: "… *not only there is substantially more correlation between absolute returns than returns themselves, but the power transformation of the absolute return also has quite high autocorrelation for long lags*" (p. 83). It can be argued that the 'long memory' feature appears to be present for which models capturing volatility need to account for.

Up and until the beginning of the 1980's the above data regularities were not addressed by the more traditional and simple time series and other models developed in order to model and forecast volatility. However based on the above mentioned problems, in 1982 Engle proposed and developed the Autoregressive Conditional Heteroscedasticity model, the ARCH model, a new methodological approach which received wide acceptance in the finance literature. This is the main topic of the next section.

2.6 ARCH/GARCH Modelling

This section describes a selection of the most popular in the finance literature models belonging to the ARCH/GARCH genre.

ARCH

Most papers refer to the test of Autoregressive Conditional Heteroscedasticity (ARCH) model. This was developed by Engle (1982), in his attempt to test for the ARCH effects on the variance of the United Kingdom inflation, and latter reviewed by Engle and Bollerslev (1986). This model accounts for the difference between the unconditional and the conditional variance of a stochastic process. While conventional econometric models operate under the assumption of a constant variance, the ARCH process allows the conditional variance to vary over time, leaving the unconditional variance constant. To model for ARCH effects in the conditional variance of a random error, ε_t we have:

$$h_t^2 = Var(\varepsilon_t \,|\, \Omega_{t-1}) \tag{2.12}$$

where h_t^2 is the conditional volatility and Ω_{t-1} is the information set.

The ARCH (q) specification is given by:

$$\mathrm{h}_t^2 = \omega + \sum_{i=1}^{a} a_i \varepsilon_{t-i}^2 \tag{2.13}$$

GARCH (Generalised ARCH)

The GARCH model of Engle (1982) and Bolerslev (1986) requires joint estimation of the conditional mean model and the variance process. On the assumption that the conditional mean stochastic error, ε_t, is normally distributed with zero mean and time-varying conditional variance, h_t^2, the GARCH (1,1) model is given by:

$$h_{t+1}^2 = \omega + \alpha\varepsilon_t^2 + \beta h_t^2 \qquad (2.14)$$

where all the parameters must be positive, while the sum of $\alpha + \beta$ quantifies the persistence of shocks to volatility. The GARCH (1,1) model generates one-step-ahead forecasts of volatility as a weighted average of the constant long-run or average variance, ω, the previous forecast variance, h_t^2, and previous volatility reflecting squared 'news' about the return, ε_t^2. In particular, as volatility forecasts are increased following a large return of either sign, the GARCH specification captures the well-known volatility clustering effect. The GARCH models are also capable of capturing leptokurtosis, skewness (besides volatility clustering), which are the features most often observed in empirical analysis.

EGARCH (nonsymmetrical dependencies)

Depending on the nature of the data the researchers are investigating, several variations of the above models are used. For example, a more appropriate technique that incorporates asymmetries in the modelling of volatility is the Exponential GARCH or EGARCH model introduced by Nelson (1991). This approach captures the skewness and allows the ARCH process to be asymmetrical. For investigation of volatility spillovers, pairwise univariate EGARCH models are used. An example is

illustrated by Appiah-Kusi and Pescetto (1998) when examining the spillover effect of the "Tequila effect" on African countries. The EGARCH model of Nelson (1991) provides an alternative asymmetric model:

$$\log(h_{t+1}^2) = \omega + \alpha\left|\frac{\varepsilon_t}{h_t}\right| + \gamma\frac{\varepsilon_t}{h_t} + \beta\log(h_t^2) \qquad (2.15)$$

where the coefficient γ captures the asymmetric impact of news with negative shocks having a greater impact than positive shocks of equal magnitude if $\gamma<0$, while the volatility clustering effect is captured by a significant α. Finally, the use of the logarithm form allows the parameters to be negative without the conditional variance becoming negative.

TGARCH (Threshold-GARCH, non-symmetrical dependencies)

The GARCH model, although non-linear in the conditional mean error postulates a linear dependence of conditional variance upon squared past errors and past variances, such that opposite shocks of equal magnitude inevitably incur the same effect upon variance. A significant issue that has arisen in the empirical application of GARCH models to financial data, and equity market data in particular, concerns the potential for an asymmetric effect of positive and negative shocks upon conditional variance. As noted by Black (1976), and expanded upon further by Christie (1982), a negative relationship often holds between current variance and the sign of past shocks. Thus, a negative shock increases the conditional variance by a greater amount than an equal positive shock, thereby generating the so-called 'leverage effect'. We therefore

consider one of the most popular asymmetric-GARCH models, namely the threshold-GARCH (TGARCH) model of Glosten, Jagannathan and Runkle (1993):

$$h_{t+1}^2 = \omega + \alpha\varepsilon_t^2 + \gamma\varepsilon_t^2 I_t + \beta h_t^2 \qquad (2.16)$$

where the leverage effect is captured by the dummy variable I_t, such that I_t =1 if $\varepsilon_{t-1}<0$, and I_t=0 if $\varepsilon_{t-1}>0$. Thus, in the TGARCH (1,1) model, positive news has an impact of α, and negative news has an impact of $\alpha+\gamma$, with negative (positive) news having a greater effect on volatility if $\gamma > 0$ ($\gamma < 0$).

APARCH (Asymmetric Power ARCH)

An alternative model capturing the information asymmetry developed by Ding et al. (1993) is the asymmetric power-ARCH, APARCH model, where the power parameter on the standard deviation is estimated and not imposed:

$$h_t^\delta = \omega + \alpha_1(|\varepsilon_{t-1}| - \gamma\varepsilon_{t-1})^\delta + \beta_1 h_{t-1}^\delta \qquad (2.17)$$

Where $\delta>0$ and γ captures any asymmetric effect of positive and negative news upon volatility.

QGARCH (Quadratic GARCH)

A model that copes with skewed returns in a similar way to the GJR model is the Quadratic GARCH model proposed by Engle and Nq (1993) and further developed by Sentana (1995):

$$h_{t+1} = \omega + \alpha(e_t - \gamma)^2 + \beta h_t \quad (2.18)$$

When γ takes a positive value, it can be seen that a negative e_t value has a greater impact on h_{t+1}.

IGARCH (Integrated GARCH and other long memory GARCH type models)[10]

Engle and Bollerslev (1986) also put forward the integrated GARCH the IGARCH an extension to GARCH model. In order to capture the characteristic of volatility persistence, the GARCH model features an exponential decay in the autocorrelation of conditional variances. It has been noted that squared and absolute returns of financial assets typically exhibit serial correlations that are slow to decay, similar to those of an integrated *I (d)* process. Shocks in the volatility series seem to have long memory and lasting impact on future volatility over a long horizon. The IGARCH captures this effect but a shock in this model impacts upon future volatility over an infinite horizon and the conditional variances does not exist for this model, Granger and Poon (2003). The IGARCH model tries to specify the second moment of a financial series and the mathematical expression of the model is similar to that of a GARCH model equation (2.12) for which the following condition must hold in the case of an IGARCH model, $\alpha + \beta = 1$ for the conditional variance to be clearly non-stationary.

Following on from the IGARCH model two more models were developed the Fractionally IGARCH, FIGARCH by Baillie, Bollerslev and Mikkelsen (1996), and

[10] See McMillan and Speight (2004).

the Fractionally Integrated Exponential GARCH, the FIEGARCH model by Bollerslev and Mikkelsen (1996). The FIGARCH model is defined as:

$$\phi(L)(1-L)^d \varepsilon_t^2 = \omega + [1-\beta(L)]\nu_t \qquad (2.19)$$

where $0 < d < 1$, such that the model in (2.18) reduces to a GARCH model for $d = 0$ and to an IGARCH model for $d = 1$. For $0 \leq d \leq 1$, the conditional variance exhibits long memory with a slow hyperbolic rate of decay from volatility shocks. The conditional variance of the FIGARCH model is given by:

$$h_t^2 = \omega + [1-\beta(L)]^{-1} + \{1-[1-\beta(L)]^{-1}\phi(L)(1-L)^d\}\varepsilon_t \qquad (2.20)$$

Davidson (2004) proposed the Hyperbolic GARCH model by generalising the FIGARCH model in order to overcome its main problem of not defining the unconditional variance:

$$h_t^2 = \omega + [1-\beta(L)]^{-1} + \{1-[1-\beta(L)]^{-1}\phi(L)[1+\tau((1-L)^d - 1)]\}\varepsilon_t \qquad (2.21)$$

The HYGARCH model nests the FIGARCH model if $\tau = 1$, and nests the GARCH model under the restriction $\tau = 0$ (or $d=0$). When $d=1$ the parameter τ becomes an autoregressive root and the HYGARCH reduces to a stationary GARCH ($\tau<1$), an IGARCH ($\tau=1$) or an explosive GARCH ($\tau>1$).

CGARCH (Component-GARCH)

The Component GARCH (CGARCH) model of Engle and Lee (1993) attempts to separate long-run and short-run volatility effects in a fashion similar to the Beveridge-Nelson (1981) decomposition of conditional mean ARMA models for economic time-series. Thus, whilst the GARCH model and its asymmetric extensions exhibit mean reversion in volatility to ω, the component GARCH model allows mean reversion to a time-varying trend, q_t. The component model specification is:

$$h_{t+1}^2 = q_{t+1} + \alpha(\varepsilon_t^2 - q_t) + \beta(h_t^2 - q_t); \quad q_{t+1} = \omega + \rho q_t + \phi(\varepsilon_t^2 - h_t^2) \qquad (2.22)$$

where q_t represents long-run (or trend) volatility provided $\rho > (\alpha + \beta)$. The forecasting error (ε_t^2 - h_t^2) serves as the driving force for the time-dependent movement of the trend, and the difference between the conditional variance and its trend (h_t^2 - q_t) is the transitory component of the conditional variance. Stationarity is achieved provided $(\alpha + \beta)(1 - \rho) + \rho < 1$, which in turn requires $\rho < 1$ and $(\alpha + \beta) < 1$. The transitory component then converges to zero with powers of $\alpha + \beta$, whilst the long-run component converges on q_t with powers of ρ.

2.7 Time varying GARCH

The finance literature has argued (Mikosh & Starica, 2004; and Terasvita 2006), that the assumption that GARCH models have constant parameters may not be appropriate when the series modelled are long. It was also documented earlier; that structural breaks in the volatility process can give rise to spurious volatility persistence if a

GARCH model is fitted to the data without accounting for breaks (Lamoureux and Lastrapes, 1990; and Mikosch and Starica, 2004).

As Terasvita (2006) mentions parameter stability is testable, an example of such a test was developed by Chu (1995), and if rejected, then the model can be generalised.

The above suggestions lead us to the conclusion that there are relevant settings in which the dynamic structure of volatility cannot be adequately captured by constant parameter GARCH models. Therefore, especially within the more recent literature a growing interest into developing adaptive volatility models, characterised by time varying parameters, allowing to account for both structural breaks as well as state dependence of the volatility response (Bauwens and Storti, 2007).

Examples of such models mentioned in Bauwens and Storti (2007), are the ST-GARCH -Smooth Transition GARCH model developed by Luukkonen et al (1988) which allows for a flexible parameterisation of the model components. The RS-GARCH –Regime Switching GARCH model by Hamilton and Susmel (1994) and Gray (1996) which allows for state dependencies in the dynamics of the volatility process. Due to practical difficulties arising from the estimation of RS-GARCH models few variations have been developed over the years Bauwens et al. (2006). A further model more recently developed in an attempt to address some of the estimation problems is the WGARCH –Weighted GARCH by Bauwens and Storti (2007), which generalises the CGARCH model, previously discussed, by Ding and Granger (1996).

There is evidence that time varying parameters models have performed better than the standard GARCH models, both in sample and out of sample (Gray 1996; and Klaassen 2002). On the other hand, regime switching GARCH models often require a considerable number of extra parameters and are often difficult to trace computationally and with little if any intuitive foundation (Frijns et al. 2011). Due to the above practical problems time varying models are not considered further in this thesis.

2.8 State of the literature

Financial market volatility has been an important research topic for the past decades and many papers have been written looking at volatility modelling and volatility forecasting in different markets.

Developments in financial econometrics over the years have suggested the use of nonlinear time series structures to model the attitude of investors toward risk and expected return. For example, Bera and Higgins (1993, p.315) mention that "*a major contribution of the ARCH literature is the finding that apparent changes in the volatility of economic time series may be predictable and result from a specific type of nonlinear dependence rather than exogenous structural changes in variables.*"

Campbell, Lo, and MacKinlay (1997, p.481) argued that "*it is both logically inconsistent and statistically inefficient to use volatility measures that are based on the assumption of constant volatility over some period when the resulting series*

moves through time." In the case of financial data, for example, large and small errors tend to occur in clusters, i.e., large returns are followed by more large returns, and small returns by more small returns. This suggests that returns are serially correlated. It can be seen that ARCH/GARCH models are strong in volatility forecasting, confirming the work of Andersen and Bollerslev (1997).

Frances and Van Dijk (1996) studied the performance of the GARCH model and two of its non-linear modifications to forecast weekly stock market volatility. The models were the Quadratic GARCH or QGARCH first introduced by Engle and Ng (1993) and the Glosten, Jagannathan and Runkle (1992) or GJR model, also known as the Threshold GARCH (TGARCH). These models describe the often observed negative skewness in stock market indices. The QGARCH model proved to be superior to the GJR model when the estimation sample does not contain extreme events such as stock market crashes.

Walsh and Yu-Gen Tsou (1998) used four methods of volatility forecasting: First, the naïve approach: this method uses past sample volatility to forecast future volatility, for an example see Alford and Boatsman (1992). Second, the Improved Extreme-Value (IEV) method by Kunitomo (1992): this model encompasses extreme observations and a drift term in the stochastic process. Third, the ARCH/GARCH models and fourth an Exponentially Weighted Moving Average (EWMA) model. The EWMA model was first used by Akgiray (1989) and involves forecasting volatility as a weighted average of previously observed volatilities. The study concluded that the EWMA technique appears to be the best forecasting technique, closely followed by

the appropriate GARCH specification. Both the IEV and historical approaches were poor by comparison.

Chong et al. (1999) in their study, use six variations of the GARCH models, the stationary GARCH, the unconstrained GARCH, the non-negative GARCH, the GARCH-M, the EGARCH and IGARCH. It was found that EGARCH was the best and IGARCH was the worst of the models.

Although, a popular research topic, the literature has not reached any conclusions regarding the 'best' model in calculating or forecasting volatility per se. A variety of volatility models have been proposed over the years from the simple standard deviation of returns and the 'simple' models such as: Random Walk, Historical Average, Simple Moving Averages, Exponential Smoothing, EWMA (Riskmetrics), Simple Regression (volatility function of its past values and an error term), Autoregressive models (ARMA, ARIMA, ARFIMA, Threshold Autoregressive). In 1982 the introduction of the autoregressive conditional heteroscedasticity model (ARCH) by Engle and its subsequent generalisation the GARCH model by Bollerslev in 1986 – was an important milestone in the volatility modelling literature. This triggered the development of other ARCH/GARCH models. The most influential models were the earlier models such as the EGARCH of Nelson (1991) and the asymmetric models of Glosten, Janagathan and Runkle (1993), Rabemananjara and Zakoian (1993), Engle and Ng (1993). More recently a further model belonging to the ARCH genre was developed, the CEV-ARCH by Fornary and Mele (2005) exploring additional properties. The list continues with more models being proposed, giving the opportunity to non-ARCH supporters, such as Figlewski, to be critical of the inception

and complexity of the ARCH type models with his remark at a conference presentation about the YAARCH model – an acronym for **Y**et **A**nother **ARCH** model, (UCSD, 1995).

In the ongoing debate of finding the best model for forecasting volatility, early empirical studies examining the forecast ability of the available models including those belonging to the GARCH family and the more simple models often concluded in favour of the simple models Cumby et al. (1993) and Jorion (1995 and 1996). Figlewski represents the academics in favour of the simple models and argued that volatility models based on simple moving averages of historical volatility are better than the more advanced GARCH type models, Figlewski (1997). Failures of the GARCH type models were also reported in the literature, Tse (1991) and Frances and Dijk (1996). On the other hand and while still no generally accepted conclusion was reached several studies produced results in favour of the GARCH type models Andersen and Bollerslev (1986), Akigiray (1989), Brailsford and Faff (1996), Andersen et al. (1999), McMillan et al. (2000) and McMillan & Speight (2004).

More recently Engle (2002) looks at the usefulness of the GARCH models over the years. In trading options for example, volatility models acted as indicators of options mispricing leading to trading opportunities. This was evident initially but more recent data has failed to support the view that ARCH volatility models lead to significant trading opportunities. As Engle (2002) mentions this is not surprising since ARCH models have a limited information set and are available to all traders today. In asset pricing too ARCH modelling played a considerable role. The theory of asset pricing is based upon the reward for bearing risk and ARCH models were developed to measure

the price of risk. The first such model was the univariate ARCH-M model of Engle, Lilien and Robins (1987). Estimation of the CAPM began with Bollerslev, Engle and Wooldridge (1988) and continued till the more recent years. Even more recently Value at Risk (VaR) analysis introduced a new role for GARCH modelling. These are some of the areas Engle mentions when looking at the GARCH literature. The main focus of his paper is not about the past achievements of the GARCH type models but what he sees as the five new frontiers of ARCH which he identifies as: High-Frequency Volatility Models; Multivariate Models; Options Pricing and Hedging; Application of GARCH models to the broad class of non-negative processes and Simulation methods for conditional expectations (use of Least Squares Monte Carlo to examine non-linear properties).

Since 2002 several of the above mentioned areas by Engle have been looked into in the finance literature, however high frequency volatility models have dominated over the rest. The availability of high frequency and intra-daily observations has introduced the concept of *realised volatility*. Volatility measures derived from high frequency data should be more accurate allowing this way for forecast efficiency gains, but Engle and Gallo (2006) state the dependence of the measure upon the frequency of observation of data makes it difficult to come to clear conclusions. One major problem identified is serial correlation in the returns. According to the same source, although the literature on realised volatility did deliver promising results still the same question remains; how can the accuracy of volatility forecasts be improved in the medium to long run?[11]

[11] A second question was also raised by Engle and Gallo (2006), should daily or intra-daily data be used in forecasting exercises?

According to Hansen et al, (2010) high frequency data and the inception of realised volatility initiated several realised measures of volatility including the realised variance, bipower variation, realised kernel and other measures described in a number of studies.[12] These measures are proven to be more informative about the current level of volatility making this way realised volatility useful in the modelling and forecasting of volatility. The estimation of GARCH models that include a realised measure in the GARCH specification were put forward by Hansen et al. (2010).

Two more important factors considered within the volatility forecasting literature are the Volatility Index (VIX) and trading volume. Since its inception, VIX has been associated and considered an important factor of volatility since the options market is a good source of information about volatility Engle (2003). On the other hand trading volume is associated with information flow and a number of studies have demonstrated that the performance of volatility models can be significantly improved with the inclusion of proxies of information flow in their model specification Taylor (2008).

The latest review paper looking at the wide variety and different types of models used in the volatility forecasting literature was by Poon and Granger (2003 and 2005). They review 93 papers on the topic and after classifying the different models into four categories; Historical Volatility models, GARCH type models, Option Implied Standard Deviation[13] models and Stochastic Volatility models they conclude the following: First, Historical Volatility models outperform the GARCH type models in 22 studies and the GARCH models are found to be better in 17 studies. Second,

[12] Studies such as Andersen et al, (2001), Barndorf-Nielsen and Shephad (2002 and 2004), Barndorf-Nielsen et al, (2008), Andersen et al, (2008) and Hansen and Horel (2009).
[13] Based on the Black-Scholes model and other variations.

Historical Volatility models were found to be better than Option Implied Standard Deviation models in 8 studies and in 26 Option Implied Standard Deviation models were found to be superior. Finally, GARCH models outperformed Option Implied Standard Deviation models in 1 study and the reverse result was found in 17 papers.

As can be seen the results are not clear cut. Overall the Option Implied Standard Deviation models appear to provide better forecasts than Historical Volatility models and GARCH type models which are ranked almost similarly. An additional important aspect of the study by Poor and Granger (2003 & 2005), is that they conclude that financial market is forecastable. The question remains in identifying the appropriate models and the relevant parameters that would assist in producing more accurate volatility forecasts.

It can be argued that in the recent literature the topic of introducing new models has shifted to the efficient estimation of the existing models. Comparisons within the GARCH class of models in studies by Hansen and Lunde (2005) suggest that the more parsimonious model the GARCH (1,1) is superior to the other models of the class after carrying out a formal test of superiority. In addition in the literature model comparisons have also been carried out within risk management frameworks for instance a Value at Risk (VaR) environment, Kuester et al. (2006), Dimitrakopoulos et al. (2010) –they more specifically look at emerging economies, and also Brownlees and Gallo (2010), who use a selection of volatility measures such as GARCH, unconditional variance, historical simulation, and RiskMetrics to find that these models are outperformed when Ultra High Frequency Data (UHFD) volatility

measures[14] are used. According to McAleer and Caporin (2011), when making comparisons a problem arises: How can comparisons and rankings take place when models are characterised by different structures? In their recent discussion paper they provide an empirical comparison of a set of models over 89 US equities, using a range of direct and indirect model comparisons taking also into account cross-sectional influences.[15] What is concluded is that more research is required on the topic focusing on the methodological approach to model comparison and robustness of model rankings.

In this thesis we will be entering the ongoing debate for determining the best model for producing the most accurate volatility forecasts. As already discussed this is a popular exercise within the finance literature, however so far no generally accepted conclusion has been reached. The forecast ability of several different types of models capturing several different attributes of the data (clustering, asymmetric dependencies and long memory) of a large dataset, will be compared. Furthermore the results are categorised on the criterion of country classification (emerging/developed) in an attempt to identify any trends or patterns. A question previously overlooked in the volatility forecasting literature, is determining the optimal in-sample period for producing out of sample forecasts. 'Backward recursion' forecasts are used in a volatility forecasting exercise to address this question. The third aspect explored in this thesis is a model comparison within a risk management framework and more specifically within a Value at Risk (VaR) setting. Here the question asked is: Are RiskMetrics forecasts good enough? To answer this question a VaR comparison exercise by exception is performed. Finally, a further two parameters are considered

[14] The realised volatility measure by Andersen et al. (2001) has become the benchmark of UHFD volatility measures.
[15] Comparing models over an increasing number of variables.

on the mission of improving volatility forecasts. The effect of the Volatility Index (VIX) and Trading Volume is assessed and a further model comparison exercise is carried out.

3. A volatility forecasting exercise

‘Simple’ versus GARCH type modes & emerging versus developed economies

Abstract

In this chapter a straightforward comparison of volatility models is performed. This has been a popular theme within the finance literature however a generally accepted conclusion has still to be reached. Entering this debate a comparison between two popular ‘simple’ models, namely the Exponential Smoothing and the Moving Averages and the more ‘advanced’ GARCH type models capturing the features of volatility clustering, the leverage effect and volatility persistence, which are found to exist in the data. Four measures of comparison are used in this exercise and a further dimension is explored based on the classification of the sample markets in order to identify the existence or not of any differences between emerging and developed economies. The results show that the more advanced GARCH type models do a better job overall than the simple models. More specifically in the order of the asymmetric models first followed by the long memory models and finally in third place the simpler time series models. When the country classification is taken into account a clearer picture emerges in the ranking of the results for the developed economies than for the developing economies, however for both the developed and emerging economies there is no contest in identifying the worst performing model.

3.1 Introduction

Stock market volatility has been the subject of numerous studies in the finance literature, particularly after the stock market crash of 1987. Likewise, modelling and forecasting volatility has also been a popular area of research within stock market volatility, for academics and practitioners alike, but with often conflicting results.

The topic of volatility forecasting is of significant importance to anyone involved in the financial markets. The magnitude of the research carried out reflects the importance of volatility in several financial and business activities. In general volatility has been associated with risk, and high volatility is thought of as a symptom of market disruption, with securities unfairly priced and the malfunctioning of the market. Especially within the derivative security market forecasting volatility is vital as managing the exposure of investment portfolios is crucial.

There are two broad categories of models used in the literature for volatility forecasting. These are: time series models with two very important sub categories the 'simpler' models and GARCH class models, and the stochastic volatility models that use market estimates from option prices. The focus of this chapter lies on the time series models of volatility forecasting.

The tendency for stock market volatility to exhibit 'clustering' has been recognised in the past (Mandelbrot, 1963; and Fama, 1965). Although all time series models capture volatility clustering it is only after the introduction of the Autoregressive Conditional Heteroscadasticity (ARCH) model by Engle (1982) and its generalisation (GARCH)

by Bollerslev (1986) and Taylor (1986) that the second and higher moments have been formally modelled. In addition, some of the models take into account volatility asymmetry.

The simplest historical price model is the Random Walk (RW) model, extending to the Historical Average (HA) model which makes use of all historical estimates, the Moving Average (MA) disregards old observations, the Exponential Smoothing (ES) method uses all historical estimates (with the more distant observations weighting less), and the Exponential Weighted Moving Average (EWMA) model which uses only recent estimates. The RiskMetrics procedure by JP Morgan uses the EWMA method. A more flexible version of the ES is the Smooth Transition Exponential Smoothing model by Taylor (2001) where the weight depends on the size and sign of the previous return. The 'simple' regression methods (autoregressive) express volatility as a function of its past values and an error term. In this category belong the ARMA type models with all the variations such as ARIMA and ARFIMA and the Threshold Autoregressive model Poon and Granger (2003).

The more sophisticated time series models, the GARCH class models do not make use of the sample standard deviations as the 'simpler' models, but formulate the conditional variance of the returns using the maximum likelihood method. The GARCH model is found to be more parsimonious than ARCH and more specifically the GARCH (1,1)[16] is found to be the most popular structure for many financial time series.

[16] GARCH (p,q) where p are number the lags of past conditional variance and q are the number of past squared returns.

The GARCH type models that allow for non-symmetrical dependencies are the EGARCH (Exponential GARCH) model by Nelson (1991), the TGARCH (Threshold GARCH) which is also know as the GJR GARCH (Glosten, Jagannathan and Runkle, 1993) model, the QGARCH (Quadratic GARCH) of Sentana (1995) and several more models that have been developed over the years. The GARCH type models that take into account the volatility persistence feature are known as the 'long memory' models. Some of the models in this sub category are the IGARCH (Integrated GARCH) by Engle and Bollerslev (1986), the FIGARCH (Fractionally IGARCH) by Baillie et al. (1996) and the FIEGARCH by Bollerslev and Mikkelsen (1996), the CGARCH (component GARCH) model of Engle and Lee (1999) and the more recently proposed model HYGARCH (Hyperbolic GARCH) by Davidson (2004), that generalises the FIGARCH model by Baillie et al. (1996). Again the list of models is extensive and only the most widely used in the literature are mentioned here.

This chapter provides a comparative evaluation of the volatility forecast ability of the first generation ARCH class model the GARCH model, the second generation ARCH class models - the asymmetric TGARCH and EGARCH models, the third generation ARCH class models - the long memory CGARCH and HYGARCH models and two 'simple' popular and representative models, the moving average and exponential smoothing models. Four different forecast evaluation techniques are used in order to find the best and worst performing models over a large sample of 25 countries, both developed and emerging.

The remainder of the chapter is structured as follows. Section 2 looks at the issues surrounding volatility in emerging stock markets. Section 3 describes the data and

methodology employed. Section 4 describes the volatility forecasting models use for this exercise. Section 5 looks at the methods used for comparing the forecast performance of the models in section 4 and reports the outcomes of the comparative forecast exercise. In section 6 a further categorisation of results is presented based on market classification and finally section 7 summarises the findings and concludes.

3.2 Volatility in emerging markets

The majority of the studies carried out investigating the topic of stock market volatility have concentrated on developed economies. However, in an attempt to draw conclusions on a more global scale this chapter will also include emerging markets in the sample.

One of the first issues explored by the literature was the nature of volatility in an attempt to gain a better understanding of emerging equity markets Bekaert and Harvey (1997), De Santis and Imrohoroglu (1997) and Aggarwal et al. (1999). For this reason the literature has concentrated on the volatile nature of emerging markets. Generally, emerging stock markets have been characterised by both high average volatility and a wide dispersion of volatility. Furthermore, both the magnitude and the range of volatility in emerging stock markets are much greater than that found in developed stock markets. Based on these essential characteristics of emerging markets empirical investigations have attempted to provide an understanding of the nature and the determinants, of both the time-series and cross-sectional behaviour of emerging market volatility, Fifield, Lonie and Power (1998).

Richards (1996) examined the proposition that emerging stock markets returns have become more volatile in recent years. The reason for this is primarily the increased scale of institutional involvement. Analysis by Richards (1996) suggests that the period between 1975 and 1992, there was no tendency for an increase in volatility while the period between 1992 and 1995 was characterised by lower volatility than in the earlier sample period (despite the increased foreign institutional investment). His findings were supported by many studies examining capital market liberalisation. Results by Spyrou and Kassimatis (1999) suggest that the nature of volatility has not changed dramatically after liberalisation and that volatility is more likely to be unaffected or reduced following liberalisation, confirming the studies of Kim and Singal (1993) and Jun (1993). On the other hand Grabel (1995) presented evidence that volatility increased following financial liberalisation. Arestis and Demetriades (1997) argue that there still is a relationship between financial liberalisation and equity market volatility. More recent studies also failed to produce a generally accepted conclusion, Kim and Singal (2000), Jayasuriya (2005), and Cunado et al. (2006).

Emerging stock markets are more sensitive to information inflows. This could be for various reasons such as the liberalisation mentioned above, or the regulatory changes of the economies and markets to i.e. foreign influences. An example of this could be the Istanbul Stock Exchange, which underwent regulatory changes in 1991 and as consequence an increase in volatility was reported Antoniou et al. (1997).

A further issue relevant to the investigation of emerging stock market volatility is the finding that stock return volatility responds asymmetrically to news. Koutmos and Booth (1995) showed that volatility in one market is correlated with price fluctuations in different markets. This phenomenon is known as the 'spill over' effect. Both domestic and foreign investors observing price changes in one market in order to develop trading strategies in another market reinforce this. An example worth mentioning is what is known as the 'Tequila effect' where in all the markets[17] studied there was a sharp increase in volatility during 1995, due to the financial crises in Mexico. The 'Tequila effect' spilled over into the African markets too, Appiah-Kusi and Pescetto (1998). On the other hand Bekaert and Harvey (1997) suggest that increased volatility is determined also by local events and Aggarwal et al. (1999) mention that changes in volatility are sudden in emerging markets.

ARCH effects have been identified in equity markets of developed economies, thus the question raised is whether ARCH effects are also present in emerging stock markets. The presence of ARCH effects in emerging equity markets is confirmed by several studies, for example, Brooks et al. (1997), Appiah-Kusi and Menyah (2003), Hassan et al. (2003), Haque et al. (2004), and Alper and Yilmaz (2004), hence allowing us to proceed with the same rationale in modelling and forecasting with the use of a selection of GARCH type models, capturing this way the features found in our large sample. Model comparisons in emerging markets have also been the topic of interest in the finance literature, for example, Gokan (2000) and Balaban et al. (2006) with often differing results. More recently Brooks (2007) examined the applicability of the APARCH model on a good mix of developing economies from five different

[17] The markets investigated were: Botswana, Egypt, Ghana, Ivory Coast, Kenya, Mauritius, Morocco, Namibia, Nigeria, South Africa, Swaziland, Uganda, Tunisia, Zambia and Zimbabwe.

regions; Latin America, Middle East, Africa, Asia and Europe. His results are mixed with different regions having different asymmetry characteristics and also a greater range of power values unlike in developed markets.

It is evident that the investigation of emerging stock market volatility so far has given a less than clear picture. Nonetheless, emerging markets cannot be excluded nor ignored since emerging markets are important for the global economic stability.

3.3 Data and methodology

3.3.1 Data

The sample is selected from a wide geographical perspective trying to identify any possible global trends (Europe, Asia, America and Australia) and includes both developed and emerging markets, since the majority of the empirical work has been carried out mainly on developed markets. In order to draw conclusions 10 out of 25 selected countries, or 40% of the sample, are emerging markets. The selected countries in alphabetical order are: Australia, Austria, Belgium, Brazil, Chile, Denmark, France, Germany, Hong Kong, India, Indonesia, Ireland, Israel, Japan, Korea, Malaysia, Netherlands, Philippines, Singapore, Spain, Sweden, Thailand, Turkey, United Kingdom and United States of America.

Table 3.1 Sample countries by geographical region				
Geographical Region/Country	Europe	Asia	America	Australia
1	Austria	Hong Kong	Brazil *	Australia
2	Belgium	India *	Chile *	
3	Denmark	Indonesia *	USA	
4	France	Israel *		
5	Germany	Japan		
6	Ireland	Korea *		
7	Netherlands	Malaysia *		
8	Spain	Philippines *		
9	Sweden	Singapore		
10	UK	Thailand *		
11		Turkey*		
Note: * Developing/emerging economy				

Table 3.2 Sample countries by geographical region and classification					
Region\ classification	Europe	Asia	America	Australia	Total
Developed	10	3	1	1	15
Developing	0	8	2	0	10
Total	10	11	3	1	25

The choice of the different indices raises the issue of comparability, always an important issue when selecting datasets. For example, two indices are not always directly comparable i.e. the USA S&P 500 is a 'selection' (capitalisation-based) index, while the Korean KOSPI is an all-shares' index.[18] The problem here is that markets differ in terms of structure and trading activity. Because of this, high-capitalisation criterion could be used when selecting indices (where possible). This selection allows both a good approximation of the total market activity as well as the inclusion of the most liquid stocks, thus removing any thin-trading considerations. On the other hand several papers use all share indices. The aim is to use indices that are important to the country concerned and to the market participants. These differ across different countries, for example for France is the FRAC-40, for Germany the DAX-30, for the UK the FTSE-100, for Japan the Nikkei-225, etc. Moreover, the selection

[18] Example from Hwang & Salmon (2004).

of any market in the sample will not only be based upon the availability or not of a specified index i.e. high-cap index which most markets have, but also upon the availability of the index during the period of the investigation.

All the data are obtained by Datastream market information service. For all the countries daily closing price data from 1 January 1990 to 31 July 2006[19] are selected and the price indices are converted to returns by the standard method of calculating the logarithmic differences[20]. The data for each country are partitioned into the in-sample estimation periods from 1 January 1990 to 31 December 1999 (10 years or 2610 observations) and out-of-sample estimation period from 1 January 2000 to 31 July 2006 (6 years and 7 months or 1716 observations). The descriptive statistics of the returns are presented in the table 3.3.

The mean and median of the returns are broadly consistent and close to zero. In general and as expected the standard deviation of the developing markets is slightly higher compared to that of the developed with Brazil appearing to be the most volatile market and Australia the least volatile. The Jarque-Bera tests for normality are consistent with the skewness and kurtosis values and normality is rejected for all series.

[19] The purpose of this exercise is to address and discuss the modeling and forecasting of volatility using a large sample for several countries in different regions of the world. Due to the large sample selected several events have occurred during which extreme values would have occurred and are part of the sample, for example the Mexican crisis in 1994, the Asian crises in 1997, the 2001 events etc. Spillover effects and contagion issues were addressed in the literature review especially in relation to emerging markets. In order account for a good representation of reality the data sample is not adjusted for outliers.

[20] Logarithmic differences of closing prices: $\mathrm{R}_{\mathrm{t}} = \log(p_t / p_{t-1})$

Table 3.3 Descriptive statistics for all sample countries

	Mean	Median	Maximum	Minimum	St. Dev	Skewness	Kurtosis	JB
Australia	0.00025	0.000140	0.06067	-0.07449	0.00763	-0.42878	8.22123	5045.2
Austria	0.00027	0.000003	0.083548	-0.10247	0.011192	-0.28333	10.6794	10685.3
Belgium	0.00023	0.000001	0.09334	-0.06295	0.010165	0.197609	9.70867	8136.78
Brazil	0.00349	0	0.693147	-0.69315	0.049941	0.766203	95.6704	1548020
Chile	0.00075	0.000004	0.089786	-0.07666	0.011586	0.273224	7.81391	4228.93
Denmark	0.00029	0.000007	0.049699	-0.06259	0.010303	-0.33357	6.03549	1740.68
France	0.00021	0	0.070023	-0.07678	0.013044	-0.10437	6.02057	1652.05
Germany	0.00018	0.000259	0.075517	-0.09881	0.0141	-0.24918	7.15158	3150.76
Hong Kong	0.00041	0	0.172471	-0.14735	0.015473	-0.03616	13.7432	20800.09
India	0.00059	0	0.166409	-0.11936	0.016496	0.033593	10.4071	9888.08
Indonesia	0.00028	0	0.131277	-0.12732	0.014873	0.222272	13.7396	20820.87
Ireland	0.00033	0.000204	0.060406	-0.07569	0.009587	-0.40619	8.24043	5067.83
Israel	0.00064	0	0.096118	-0.11723	0.014693	-0.41177	8.98972	6587.51
Japan	-0.0002	0	0.124303	-0.07234	0.014446	0.17036	6.40987	2116.24
Korea	0.000008	0	0.100238	-0.12805	0.018764	-0.04566	7.03412	2934.24
Malaysia	0.00011	0	0.208174	-0.24153	0.014925	0.488913	45.9892	333211
Netherlands	0.00027	0.000383	0.095169	-0.07531	0.012709	-0.14144	8.26480	5009.46
Philippines	0.00017	0	0.161776	-0.09744	0.015305	0.508535	11.9252	14541.74
Singapore	0.00017	0	0.148685	-0.09672	0.012543	0.18958	14.2986	23031.27
Spain	0.00031	0.000217	0.068372	-0.08876	0.012722	-0.21955	6.69395	2493.74
Sweden	0.00035	0.000005	0.110228	-0.08527	0.014383	0.178738	6.91763	2788.83
Thailand	-0.000005	0	0.113495	-0.10028	0.017229	0.232464	7.95217	4458.38
Turkey	0.00171	0.000341	0.177736	-0.19979	0.029574	-0.06473	6.59419	2330.99
UK	0.00020	0.000002	0.059026	-0.05885	0.010069	-0.10881	6.29946	1970.36
USA	0.00029	0.000105	0.055732	-0.07113	0.009879	-0.09624	7.03941	2947.10

3.3.2 Methodology

Below is a selection of models belonging to the GARCH genre. Different models capture different features found in empirical investigations of financial market returns such as volatility clustering, information asymmetry and long memory. The focus of this exercise is mainly on the information asymmetry and long memory elements captured by the different models described in this section and the categorisation is based on this criterion. The models considered include the GARCH model, the first generation of ARCH models the Generalised ARCH symmetric model, the second generation asymmetric models the TGARCH by Glosten, Jagannathan and Runkle (1993) and the EGARCH by Nelson (1991), and the third generation long-memory models the CGARCH by Engle and Lee (1999) and the HYGARCH by Davidson (2004). In addition to the GARCH genre of models the forecasting performance of the more ‘simple’ models is also considered such as Moving Average and Exponential

Smoothing. Previous studies comparing the forecast ability of the more 'simple' models in contrast to those of the GARCH models have given mixed results dividing the opinions of those involved.

In the first instance prices are converted to returns by the standard method of calculating the logarithmic differences[21] and then in order to establish the notation and methods to be used, the returns process r_t is given by:

$$r_t = m_t + \varepsilon_t \tag{3.1}$$

where m_t is the conditional mean process, which could include autoregressive (AR) or moving average (MA) terms, and the error term can be decomposed as:

$$\varepsilon_t = \sigma_t \mathrm{z}_t \tag{3.2}$$

with z_t an idiosyncratic zero-mean and constant variance noise term, and σ_t is the volatility process to be estimated and forecast, with forecast values denoted h_t^2.

The sample data is split between the in-sample period, $t=1,\ldots,T$, and the out-of-sample period $t=T,\ldots,\tau$. To derive the 'actual volatility' series, on the basis of which volatility forecasts are compared to, we follow the Pagan and Schwert (1990) methodology in representing past volatility by the squared residuals from a

[21] Returns are calculated as $\mathrm{R}_t = \log(p_t / p_{t-1})$.

conditional mean model, such as (3.1), for returns estimated over the whole sample period.

Generalised Autoregressive Conditional Heteroscedasticity (GARCH)

The GARCH model of Engle (1982) and Bolerslev (1986) requires joint estimation of the conditional mean model (3.1) and the variance process. On the assumption that the conditional mean stochastic error, ε_t, is normally distributed with zero mean and time-varying conditional variance, h_t^2, the GARCH(1,1) model is given by:

$$h_{t+1}^2 = \omega + \alpha\varepsilon_t^2 + \beta h_t^2 \qquad (3.3)$$

where all the parameters must satisfy the non-negativity constraints $\omega > 0$ and $\alpha, \beta \geq 0$ while the sum of $\alpha + \beta$ quantifies the persistence of shocks to volatility. The GARCH (1,1) model generates one-step-ahead forecasts of volatility as a weighted average of the constant long-run or average variance, ω, the previous forecast variance, h_t^2, and previous volatility reflecting squared 'news' about the return, ε_t^2. In particular, as volatility forecasts increase following a large return of either sign, the GARCH specification captures the well-known volatility clustering effect.

Threshold-GARCH (TGARCH)

The GARCH model, although non-linear in the conditional mean error postulates a linear dependence of conditional variance upon squared past errors and past variances, such that opposite shocks of equal magnitude inevitably confer the same effect upon

variance. A significant issue that has arisen in the empirical application of GARCH models to financial data, and equity market data in particular, concerns the potential for an asymmetric effect between positive and negative shocks upon conditional variance. As noted by Black (1976), and expanded upon further by Christie (1982), a negative relationship often holds between current variance and the sign of past shocks. Thus, a negative shock increases the conditional variance by a greater amount than an equal positive shock, so generating what is known as the 'leverage effect'. We therefore consider one of the more popular asymmetric-GARCH models, namely the threshold-GARCH (TGARCH) model of Glosten, Jagannathan and Runkle (1993):

$$h_{t+1}^2 = \omega + \alpha\varepsilon_t^2 + \gamma\varepsilon_t^2 I_t + \beta h_t^2 \tag{3.4}$$

where the leverage effect is captured by the dummy variable I_t, such that $I_t = 1$ if $\varepsilon_{t-1} < 0$, and $I_t = 0$ if $\varepsilon_{t-1} > 0$. Thus, for the TGARCH(1,1) model, positive news have an impact of α, and negative news have an impact of $\alpha + \gamma$, with negative (positive) news having a greater (smaller) effect on volatility if $\gamma > 0$ ($\gamma < 0$).

Exponential-GARCH (EGARCH)

The EGARCH model of Nelson (1991) provides an alternative asymmetric model:

$$\log(h_{t+1}^2) = \omega + \alpha\left|\frac{\varepsilon_t}{h_t}\right| + \gamma\frac{\varepsilon_t}{h_t} + \beta\log(h_t^2) \tag{3.5}$$

where the coefficient γ captures the asymmetric impact of news with negative shocks having a greater impact than positive shocks of equal magnitude if $\gamma < 0$, while the

volatility clustering effect is captured by a significant α. Finally, the use of the logarithmic form allows the parameters to be negative without the conditional variance becoming negative.

Component-GARCH (CGARCH)

The component GARCH (CGARCH) model by Engle and Lee (1999) attempts to separate long-run and short-run volatility effects in a fashion similar to the Beveridge-Nelson (1981) decomposition of conditional mean ARMA models for economic time-series. Thus, whilst the GARCH model and its asymmetric extensions exhibit mean reversion in volatility to ω, the component GARCH model allows mean reversion to a time-varying trend, q_t The component model specification is:

$$h_{t+1}^2 = q_{t+1} + \alpha(\varepsilon_t^2 - q_t) + \beta(h_t^2 - q_t) \tag{3.6}$$

where $q_{t+1} = \omega + \rho q_t + \phi(\varepsilon_t^2 - h_t^2)$

represents long-run (or trend) volatility provided $\rho > (\alpha + \beta)$. The forecasting error $(\varepsilon_t^2 - h_t^2)$ serves as the driving force for the time-dependent movement of the trend, and the difference between the conditional variance and its trend $(h_t^2 - q_t)$ is the transitory component of the conditional variance. Stationarity is achieved provided $(\alpha + \beta)(1 - \rho) + \rho < 1$, which in turn requires $\rho < 1$ and $(\alpha + \beta) < 1$. The transitory component then converges to zero with powers of $\alpha + \beta$, whilst the long-run component converges on q_t with powers of ρ.

Hyperbolic-GARCH (HYGARCH)

This is a more recently proposed model by Davidson (2004) in an attempt to capture the characteristic of long memory. The HYGARCH model generalises the FIGARCH[22] model by Baillie et al. (1996). An alternative to the CGARCH model for long-memory, already described above, has been provided by the fractionally integrated FIGARCH-type model. In the GARCH model of (3.3), where $\alpha + \beta = 1$, the process is said to be integrated in volatility such that current and past shocks persist indefinitely in conditioning future variance (Engle and Bollerslev, 1986). We can see this using the 'ARMA-in-squares' form, where $v_t = \varepsilon_t^2 - h_t^2$ and substituting for the variances in (3.3) and rearranging we have:

$$\varepsilon_t^2 = \omega + (\alpha_1 + \beta_1)\varepsilon_{t-1}^2 + v_t - \beta_1 v_{t-1} = [1 - \alpha(L) - \beta(L)]\varepsilon_t^2 = \omega + [1 - \beta(L)]v_t \qquad (3.7)$$

Thus, squared errors follow a heteroscedastic ARMA (1,1) process. Should the autoregressive lag polynomial $[1 - \alpha(L) - \beta(L)]$ contain a unit root, the process is defined to be integrated in variance and is given by:

$$\phi(L)(1 - L)\varepsilon_t^2 = \omega + [1 - \beta(L)]v_t \qquad (3.8)$$

However, the general belief is that while volatility shocks may take a long time to decay, they nevertheless ultimately do decay (Ding et al. 1993). This prompted the FIGARCH model of Baillie et al. (1996) which is defined as:

[22] The FIGARCH model constitutes an alternative to the GARCH and IGARCH (Integrated GARCH) models by adding a fractionally integrated parameter in order to capture the long memory element.

$$\phi(L)(1-L)^d \varepsilon_t^2 = \omega + [1-\beta(L)]\nu_t \tag{3.9}$$

where $0 < d < 1$, such that this model reduces to a GARCH model for $d = 0$ and to an Integrated-GARCH model for $d = 1$. For $0 \le d \le 1$, the conditional variance exhibits long memory with a slow hyperbolic rate of decay from volatility shocks. The conditional variance of the FIGARCH model is given by:

$$h_t^2 = \omega + [1-\beta(L)]^{-1} + \{1-[1-\beta(L)]^{-1}\phi(L)(1-L)^d\}\varepsilon_t \tag{3.10}$$

However, Davidson (2004) notes, counterintuitively, that as d approaches zero the memory of the process is increasing, and that the FIGARCH model in fact belongs to the same 'knife-edge non-stationarity' class represented by the IGARCH model. Davidson (2004) thus generalises the FIGARCH model as such, which he refers to as the hyperbolic-GARCH (HYGARCH) model:

$$h_t^2 = \omega + [1-\beta(L)]^{-1} + \{1-[1-\beta(L)]^{-1}\phi(L)[1+\alpha((1-L)^d - 1)]\}\varepsilon_t \tag{3.11}$$

Thus the HYGARCH model nests the FIGARCH model if $\alpha = 1$, see Davidson (2004) for more details.[23]

Note: the sample is tested for ARCH effects and autocorrelation after the above GARCH type models were applied. With the exception of the EGARCH model for

[23] The FIGARCH model has been further criticised from a forecast modelling perspective because the unconditional variance of the FIGARCH model does not exist and its long memory characteristics are dependent upon certain second-order stationarity conditions (see Giraitis et al., 2000; and Davidson, 2004) and thus is not considered further here.

which in five countries the null hypothesis of no ARCH effects is rejected all other countries of the sample and for all remaining models the presence of ARCH effects is captured by the specified number of lags. This fining does not come as a surprise as the finance literature supports the more parsimonious GARCH type models. This is the methodology used throughout this thesis. In appendix 1 a table summing up all the results of the ARCH test is presented. Addressing the problem of Autocorrelation the "Ljung Box" test is used for one lag.[24] The standardised residuals are examined for autocorrelation. If there is no serial correlation in the residuals, the Autocorrelation and Partial Autocorrelations at all lags should be nearly zero and all Q-statistics should be insignificant with large p-values. The table presented in appendix 2 shows the results for the Ljung Box autocorrelation test. For only two sample countries when the EGARCH model is used autocorrelation is still present.

Simple models

The term 'simple' for the models described below refers to the traditional and widely used techniques not only in finance but other disciplines too. Although the list of those models is extensive, only two models will be examined. Practitioners tend to have a preference in using these models.

Moving Average (MA)

Under the moving average method, volatility is forecast by an unweighted average of previously observed volatilities over a particular historical time interval of fixed length:

[24] This is the methodology followed by Engle (2001), in his work titled: GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics. In this paper 15 lags are used.

$$h_{t+1}^2 = \frac{1}{p}\sum_{j=1}^{P}\sigma_j^2 \qquad (3.12)$$

where P is the moving average period or 'rolling window'. The choice of this interval is essentially arbitrary, however, the length chosen here equates to a window of 60 days.[25]

Exponential Smoothing (ES)

Under exponential smoothing the one-step ahead volatility forecast is a weighted function of the immediately preceding volatility forecast and actual volatility:

$$h_{t+1}^2 = \phi h_t^2 + (1-\phi)\sigma_t^2 \qquad (3.13)$$

where ϕ is a smoothing parameter constrained to lie between zero and one, such that for $\phi=0$, the exponential smoothing model reduces to a random walk model, while for $\phi=1$ weight is given only to the prior period forecast. The value of ϕ is determined empirically by that value which minimises the in-sample sum of squared prediction errors.

Exponential smoothing has been shown to be a strong model in terms of accuracy and simplicity (Poon and Granger, 2003). The simple exponential smoothing method can be viewed as a special case of the IGARCH, which is a non-stationary version of GARCH.

[25] The 60 window is approximately 3 months of trading often regarded as the length of time over which practitioners evaluate their models.

Comparisons of forecast performance

There are many methods for evaluating and comparing the accuracy of the different forecasting models used. Several of these models are reviewed in the forecasting literature for example in Diebold and Lopez (1996) and Granger and Poon (2003, 2005). Two important categories of measures for forecast comparisons are selected. In the first category the Mean Error (ME) and Mean Absolute Error are selected and in the second regression based efficiency tests are implemented comparing the coefficient of determination (R^2) from the different models. The testing procedure of Mincer-Zarnowitz (MZ) is used where the true volatility value is regressed on the forecast value. As a measure of 'true' (or 'actual') volatility against which the forecast performance of volatility is compared to, the Pagan and Schwert (1990) model is followed in using the squared error term from a conditional mean model for returns estimated over the full data set including both the in-sample and out-of-sample periods.

Mean Error (ME) and Mean Absolute Error (MAE)

The ME statistic is used as a guide of direction of over or under prediction on average whereas the MAE statistic measures the average absolute forecast error, which does not permit the offsetting effects of over-prediction and under-prediction as happens in the estimation of the ME. When the forecasting techniques are compared a lower ME and MAE are preferred.

$$ME = \frac{1}{t}\sum_{t=T+1}^{T+t}(h_t^2 - \sigma_t^2) \qquad (3.14)$$

$$MAE = \frac{1}{t}\sum_{t=T+1}^{T+t}\left|h_t^2 - \sigma_t^2\right| \tag{3.15}$$

3.4 Empirical results and analysis

Table 3.4 reports the ME statistics for the selected volatility forecast models of the 25 countries of the sample in alphabetical order. It can be seen that HYGARCH performs better than the rest of the models – in 11 out of 25 countries – producing the lowest ME value. The MA method has a minimum ME in nine countries whereas the Exponential Smoothing model has a minimum ME in four countries and for one country each the EGARCH and TGARCH methods and no minimum ME values are reported for the GARCH and CGARCH methods.

The second best models are found after working out the second minimum ME value in each row of the table (for each country). The MA model comes first this time with 7 second minimum values followed by the CGARCH with 6 second minimum values. In third place is the Exponential Smoothing model with 5 and then the EGARCH and HYGARCH with 2 second minimum values each and finally the GARCH and TGARCH models with 1 second minimum value respectively.

Likewise, the third best models are found after working out the third minimum ME value in each row of the table. As expected, the results are more dispersed, with the MA being first in six countries, second the GARCH, EGARCH, TGARCH and CGARCH models in four countries respectively, the HYGARCH in two and the Exponential Smoothing model in one country.

Finally in Table 3.4 the worst performing models are also reported, by working out the maximum ME value for each country. Here, the worst performing model appears to be the Exponential Smoothing model in 11 out of 25 countries.

In an attempt to look at the overall performance of the models when using the ME statistic, it can be seen that although the best model appears to be one of the long memory GARCH class models the HYGARCH, the 'simpler' models perform better overall - in 13 countries out of 25. The CGARCH has no minimum values consequently bringing the long memory models in the second place. The asymmetric GARCH class models (EGARCH and TGARCH) also known as asymmetric models which come in third place based on their overall performance.

As noted above a drawback of the ME statistic is that it tends to offset positive and negative forecast errors. Thus, we proceed with the MAE measure.

Table 3.4 ME statistic for volatility forecast models for all sample

Country/ Model	GARCH	EGARCH	TGARCH	CGARCH	HYGARCH	Ex. Sm	MA
Australia	-0.0000097	-0.0000073‡	-0.0000107	-0.0000079†‡	0.0000172•	-0.0000117	-0.0000022†
Austria	-0.0000164	-0.0000086†‡	-0.0000113	-0.0000154	-0.0000007†	-0.0000203•	-0.0000069‡
Belgium	0.0000199	0.0000333•	0.0000186	0.0000185	0.0000127‡	0.0000112†	0.0000143†‡
Brazil	-0.0007886	-0.0008597•	-0.0007929	-0.0007535†‡	0.0000175†	-0.0008050	-0.0002151‡
Chile	-0.0000235	-0.0000259	-0.0000244	-0.0000175‡	-0.0000105†	-0.0000533•	-0.0000195†‡
Denmark	0.0000181	0.0000216	0.0000176†‡	0.0000177	0.0000292•	0.0000091‡	0.0000089†
France	0.0000295	0.0000265†‡	0.0000283	0.0000282	0.0000551•	0.0000061†	0.0000153‡
Germany	0.0000341	0.0000403•	0.0000371	0.0000275	0.0000201†‡	0.0000196‡	0.0000189†
Hong Kong	-0.0000232	-0.0000341	-0.0000327	-0.0000182†‡	0.0000182‡	-0.0001111•	-0.0000145†
India	-0.0000179	-0.0000328•	-0.0000144	-0.0000185	-0.0000004†	-0.0000097†‡	0.0000044‡
Indonesia	-0.0000370	-0.0000393	-0.0000372	-0.0000101‡	-0.0000291†‡	-0.0000905•	-0.0000042†
Ireland	0.0000058	0.0000039	0.0000027†‡	0.0000054	0.0000184•	-0.0000007†	-0.0000007†
Israel	-0.0000301†‡	-0.0000380	-0.0000395	-0.0000327	0.0000138†	-0.0000431•	-0.0000195‡
Japan	-0.0000077†‡	-0.0000090	-0.0000116	-0.0000070‡	0.0000318•	-0.0000181	-0.0000013†
Korea	-0.0000050‡	-0.0000096	-0.0000181	-0.0000114	-0.0000030†	-0.0001049•	-0.0000067†‡
Malaysia	-0.0000201	-0.0000188†‡	-0.0000225	-0.0000158‡	-0.0000127†	-0.0001871•	-0.0000586
Netherlands	0.0000248	0.0000338•	0.0000232†‡	0.0000254	0.0000247	0.0000206‡	0.0000185†
Philippines	-0.0000244	-0.0000142†‡	-0.0000186	-0.0000235	-0.0000063†	-0.0000670•	-0.0000111‡
Singapore	-0.0000132	-0.0000176	-0.0000164	-0.0000073†‡	0.0000015†	-0.0000635•	-0.0000065‡
Spain	0.0000121	0.0000097‡	0.0000116†‡	0.0000119	0.0000512•	-0.0000178	0.0000063†
Sweden	0.0000159	0.0000238	0.0000089‡	0.0000172	0.0000414•	0.0000058†	0.0000111†‡
Thailand	-0.0000316†‡	-0.0000404	-0.0000343	-0.0000239‡	-0.0000031†	-0.0001232•	-0.0000353
Turkey	-0.0000761	-0.0000491†	-0.0000790	-0.0000583‡	0.0001050	-0.0002102•	-0.0000654†‡
UK	0.0000094	0.0000122•	0.0000014†	0.0000095	0.0000112	0.0000045‡	0.0000076†‡
USA	0.0000051†‡	0.0000085•	0.0000055	0.0000054	-0.0000008†	0.0000011‡	0.0000074

Notes: †: Best performing model in the row, ‡: Second best performing model in the row,
†‡: Third best performing model in the row, •:Worst performing model in the row.

Similarly as above, Table 3.5 reports the MAE statistics for the sample volatility forecast models of the 25 countries in alphabetical order. Unlike when using the ME statistic, the MAE statistic indicates a stronger HYGARCH model – in 16 out of 25 countries with a minimum MAE value – even though before the HYGARCH model was the worst performer. The second best model is the EGARCH model with five minimum MAE values followed by the MA with three minimum values and the CGARCH with one minimum value.

The second best models are the MA in 8 countries, the EGARCH in six countries, the TGARCH in five countries, the CGARCH in four countries and the HYGARCH in two countries. The GARCH and Exponential Smoothing models are found to have no countries where they perform as best or second best.

The third best model measure again gives us a wider dispersion. The CGARCH is better in eight countries, the GARCH and EGARCH in five respectively, the TGARCH in four, HYGARCH in two and the MA model in one country. Again, the Exponential Smoothing model shows no performance.

As can be seen the Exponential Smoothing technique once again appears to be the worst performing model with the highest MAE value in 24 out of 25 countries.

The overall conclusion of Table 3.5 is similar to that of Table 3.4. The 'long memory' GARCH models perform better but it has to be mentioned that this is due to the HYGARCH model alone because the CGARCH model which belongs in the same category appears to be neutral. However this time the results come mainly from the HYGARCH as the best and from the Exponential Smoothing as the worst performers. In second place comes the GARCH model and in third the MA 'simpler' model. The asymmetric GARCH class models appear to be impartial.

Table 3.5 MAE statistic for volatility forecast models for all sample

Country/ Model	GARCH	EGARCH	TGARCH	CGARCH	HYGARCH	Ex. Sm	MA
Australia	0.0000543	0.0000520†‡	0.0000541	0.0000532	0.0000433†	0.0000583•	0.0000508‡
Austria	0.0001009	0.0000970†‡	0.0000984	0.0001010	0.0000948†	0.0001078•	0.0000970‡
Belgium	0.0001375†‡	0.0001328†	0.0001351‡	0.0001376	0.0001410	0.0001699•	0.0001450
Brazil	0.0008917	0.0009423•	0.0008950	0.0008617†‡	0.0003320†	0.0009149	0.0004463‡
Chile	0.0000888†‡	0.0000898	0.0000892	0.0000861†	0.0000901	0.0001106•	0.0000876‡
Denmark	0.0001351	0.0001335‡	0.0001347†‡	0.0001355	0.0001313†	0.0001490•	0.0001412
France	0.0002030	0.0002014‡	0.0002015†‡	0.0002038	0.0001937†	0.0002439•	0.0002083
Germany	0.0002544	0.0002497†	0.0002508‡	0.0002544†‡	0.0002583	0.0003031•	0.0002615
Hong Kong	0.0001973	0.0001999	0.0002011	0.0001950†‡	0.0001794†	0.0002532•	0.0001896‡
India	0.0002894	0.0002992	0.0002884‡	0.0002885†‡	0.0002812†	0.0003260•	0.0002904
Indonesia	0.0002247	0.0002256	0.0002248	0.0002138‡	0.0002245†‡	0.0002593•	0.0002091†
Ireland	0.0001141†‡	0.0001137‡	0.0001150	0.0001148	0.0001080†	0.0001237•	0.0001156
Israel	0.0001813	0.0001843	0.0001864	0.0001831†‡	0.0001629†	0.0001927•	0.0001765‡
Japan	0.0002088	0.0002060†‡	0.0002089	0.0002089	0.0001960†	0.0002176•	0.0002032‡
Korea	0.0003805†‡	0.0003809	0.0003847	0.0003819	0.0003784‡	0.0004454•	0.0003775†
Malaysia	0.0001027	0.0001023†‡	0.0001048	0.0000991‡	0.0000964†	0.0002342•	0.0001313
Netherlands	0.0002231	0.0002193†	0.0002210‡	0.0002229†‡	0.0002236	0.0002779•	0.0002341
Philippines	0.0002196	0.0002085‡	0.0002136	0.0002194	0.0002092†‡	0.0002420•	0.0002066†
Singapore	0.0001446	0.0001440	0.0001461	0.0001421†‡	0.0001381†	0.0001794•	0.0001408‡
Spain	0.0001760	0.0001743‡	0.0001746†‡	0.0001770	0.0001631†	0.0002139•	0.0001801
Sweden	0.0002581	0.0002517‡	0.0002572†‡	0.0002579	0.0002493†	0.0002947•	0.0002593
Thailand	0.0002204†‡	0.0002252	0.0002210	0.0002163‡	0.0002066†	0.0002823•	0.0002253
Turkey	0.0008415	0.0008288†‡	0.0008423	0.0008248‡	0.0007568†	0.0009784•	0.0008424
UK	0.0001317	0.0001297†	0.0001332	0.0001315†‡	0.0001302‡	0.0001562•	0.0001365
USA	0.0001306	0.0001265†	0.0001275‡	0.0001306	0.0001319	0.0001464•	0.0001297†‡

Notes: †: Best performing model in the row, ‡: Second best performing model in the row, †‡: Third best performing model in the row, •:Worst performing model in the row.

Regression Analysis (Mincer-Zarnowitz, MZ test)

The testing procedure of Mincer-Zarnowitz (1969), hereafter MZ test, is used where the true volatility value is regressed on the forecast value and the coefficient of determination is obtained R^2 for comparison purposes. However, one drawback of the MZ test is that large values have a larger impact on the regression results, thus in order to deal with this problem the same general form of the MZ regression is adopted but logarithms are used to rescale the parameters, a solution proposed by Pagan and Schwert (1990). The two regressions are:

$$\sigma_t^2 = \alpha + \beta h_t^2 + \varepsilon_t \tag{3.16}$$

$$\log(\sigma_t^2) = \alpha + \beta \log(h_t^2) + \varepsilon \tag{3.17}$$

The coefficient of determination R^2 is reported for all the regressions, representing the information content of the particular model used. In this case a higher R^2 value will be preferred.

Tables 3.6 and 3.7 report the R^2 statistic for all the regressions run comparing the selected volatility forecast models of the 25 countries of the sample in alphabetical order.

In Table 3.6 the R^2's of the first regression described above are reported. In 16 out of 25 countries the EGARCH model outperforms the rest of the models. MA has a higher R^2 in four countries, the CGARCH and HYGARCH in two countries respectively and the TGARCH in one country. No higher R^2 values are recorded for the GARCH and Exponential Smoothing models.

Looking at the second best performing model by estimating the second maximum R^2 value the TGARCH model outperforms the rest of the models in 14 out 25 countries. Second is the CGARCH model in four countries, the EGARCH and HYGARCH model are in three countries respectively and the GARCH model in one

country. There are no second higher R^2 values reported for the Exponential Smoothing once more and also this time for the MA technique.

The third best performing model gives again no result for the Exponential Smoothing model, but does for the rest of the models. The HYGARCH model is third best in eight countries, the GARCH model in seven, the TGARCH in five, two countries for the CGARCH and MA techniques and one for EGARCH.

Trying to identify the poorest performing model, not surprisingly is the Exponential Smoothing model, in 24 out 25 countries. Only in one country is the CGARCH model the worst.

Note: A further measure of forecasting performance the Mean Absolute Percentage Error (MAPE) is employed in addition to the ME and MAE. The attractive property of the MAPE is that it can be interpreted as a percentage error, and furthermore its value is bounded from below by zero. The results which can be found in appendix 3 do not give a different picture from what was previously found. The worst performing model is the Exponential Smoothing model and the best performing model is the HYGARCH models followed by the EGARCH, MA and CGARCH models.

Table 3.6 R^2 statistic for volatility forecast models of all sample countries

Country/ Model	GARCH	EGARCH	TGARCH	CGARCH	HYGARCH	Ex. Sm	MA
Australia	0.0439600†‡	0.0661530†	0.0538110‡	0.0391820	0.0396710	0.0044500•	0.0362370
Austria	0.0692100†‡	0.0753680†	0.0743310‡	0.0563280	0.0507940	0.0000530•	0.0416130
Belgium	0.1716340	0.2082810†	0.1965110‡	0.1828970†‡	0.1800390	0.0008600•	0.1149950
Brazil	0.0397870†‡	0.0506040†	0.0407940‡	0.0355280	0.0236090	0.0135010•	0.0386320
Chile	0.0631800†‡	0.0677110†	0.0644600‡	0.0627150	0.0078830	0.0075790•	0.0457310
Denmark	0.1209240	0.1397390†	0.1296420‡	0.1180600	0.1277720†‡	0.0000050•	0.0756260
France	0.1936450†‡	0.2230550†	0.2144810‡	0.1888890	0.1911320	0.0039100•	0.1595550
Germany	0.2079010	0.2360320†	0.2298220‡	0.2194850†‡	0.2111900	0.0031690•	0.1823330
Hong Kong	0.0508540	0.0710980‡	0.0538980†‡	0.0481780•	0.0497050	0.0492820	0.0752050†
India	0.1167570†‡	0.0969080	0.1132260	0.1320490†	0.1255220	0.0038220•	0.0773180
Indonesia	0.0197210‡	0.0191530†‡	0.0199410†	0.0131210	0.0155960	0.0015470•	0.0159440
Ireland	0.0593430	0.0719660†	0.0645930†‡	0.0570530	0.0695980‡	0.0016650•	0.0490630
Israel	0.0379870	0.0488220†	0.0394160	0.0409500‡	0.0401470†‡	0.0147290•	0.0358390
Japan	0.0419220	0.0488600†	0.0432740‡	0.0391380	0.0405330	0.0082910•	0.0422390
Korea	0.0469650	0.0511460	0.0500490	0.0529290‡	0.0518960†‡	0.0343120•	0.0710230†
Malaysia	0.0611150	0.0591080	0.0576390	0.0689130‡	0.0912480†	0.0539510•	0.0630980†‡
Netherlands	0.2322840	0.2538080†	0.2514960‡	0.2290040	0.2338590†‡	0.0036100•	0.1665290
Philippines	0.0020900	0.0044520‡	0.0040200†‡	0.0025430	0.0020700	0.0018680•	0.0315260†
Singapore	0.0379210†‡	0.0474490‡	0.0374970	0.0298240	0.0336720	0.0294580•	0.0486200†
Spain	0.1730530	0.1993080†	0.1915030‡	0.1638660	0.1745510	0.0208830•	0.1466670
Sweden	0.1151830	0.1460580†	0.1404430‡	0.1139850	0.1156630	0.0027630•	0.1105530
Thailand	0.0835390	0.0775040	0.0934750†‡	0.0982880‡	0.1021290†	0.0336620•	0.0596080
Turkey	0.1165890	0.1127900	0.1201330†‡	0.1301860†	0.1255700‡	0.0240740•	0.0921830
UK	0.1922480	0.2126220†	0.2108590‡	0.1952620	0.2025500†‡	0.0029040•	0.1404120
USA	0.1257020	0.1807250†	0.1699450‡	0.1237780	0.1349250†‡	0.0031070•	0.1254140

Notes: †: Best performing model in the row, ‡: Second best performing model in the row,
†‡: Third best performing model in the row, •:Worst performing model in the row.

Here the overall results suggest that the GARCH type models are by far better outperforming the simple models. More specifically, the asymmetric models perform better than the long memory models, which in turn perform better than the simple models.

In Table 3.7 the GARCH type models of asymmetric information appear to be superior as in nine countries for the EGARCH model and in eight countries for the TGARCH model the highest R^2 is recorded. In the third place the MA model has the

maximum value for five countries, the HYGARCH in two countries and the GARCH model in one country. The CGARCH and Exponential Smoothing models show on maximum value in any country.

Looking at the second best performers the picture does not change as the EGARCH and TGARCH are better in eight countries each, the HYGARCH in three countries, the GARCH and MA models in two countries and the CGARCH and Exponential Smoothing in one country each.

The third best performers in each row are the GARCH, CGARCH, HYGARCH and MA in five countries respectively. TGARCH performs better in four countries and the EGARCH model in one. No countries have a third higher value for the Exponential Smoothing model, which appears once more to be the worst performing model. In 22 out 25 countries the Exponential Smoothing model gives the minimum R^2 values.

Overall the GARCH type models have a better performance with EGARCH and TGARCH being the top techniques. The MA technique however also performs rather well. Grouping the results into the categorisation used in this exercise the asymmetric GARCH models are superior, the simple GARCH model comes second and the results for the long memory and simple models are divided.

Table 3.7 LOG R^2 statistic for volatility forecast models of all sample countries

Country/ Model	GARCH	EGARCH	TGARCH	CGARCH	HYGARCH	Ex. Sm	MA
Australia	0.0455680	0.0685900†	0.0623100‡	0.0444730	0.0462480	0.0034460•	0.0493230†‡
Austria	0.0398980†‡	0.0407030‡	0.0420640†	0.0369490	0.0286140	0.0004220•	0.0368900
Belgium	0.1353830†‡	0.1417450‡	0.1419380†	0.1344420	0.1323460	0.0009860•	0.1268330
Brazil	0.0218210†‡	0.0250380†	0.0222150‡	0.0181770	0.0106530•	0.0160050	0.0211340
Chile	0.0294800	0.0320360†	0.0304210‡	0.0301720†‡	0.0015920	0.0009440•	0.0230740
Denmark	0.0754470	0.0812940†	0.0754920	0.0784430†‡	0.0802060‡	0.0003080•	0.0639220
France	0.1083070	0.1283050†	0.1211710‡	0.1066600	0.1152120	0.0000100•	0.1158010†‡
Germany	0.1685250†‡	0.1806310†	0.1782230‡	0.1671110	0.1670500	0.0004260•	0.1644680
Hong Kong	0.0695530	0.0796490†‡	0.0709480	0.0729430	0.0752970	0.0807640•	0.0973260†
India	0.0864300†	0.0799710	0.0852160‡	0.0840990†‡	0.0836900	0.0033960•	0.0717610
Indonesia	0.0138490‡	0.0098140	0.0140020†	0.0107730	0.0125280†‡	0.0041960	0.0031270
Ireland	0.0512700	0.0573740†	0.0533990	0.0495920	0.0557620	0.0065380•	0.0542800†‡
Israel	0.0367300	0.0383280‡	0.0397440†	0.0369690	0.0381140†‡	0.0158350•	0.0298770
Japan	0.0292010	0.0430400‡	0.0382070†‡	0.0271850	0.0305690	0.0116520•	0.0434610†
Korea	0.0600380	0.0624760	0.0648960‡	0.0612530	0.0625900†‡	0.0453010•	0.0725860†
Malaysia	0.0881460	0.0945220	0.0871060•	0.0960610†‡	0.1041460	0.0935900	0.1043990†
Netherlands	0.1663450	0.1710200‡	0.1730010†	0.1669700†‡	0.1667370	0.0005960•	0.1545450
Philippines	0.0151090†‡	0.0140630	0.0156820	0.0148710	0.0161840†	0.0000720•	0.0102600
Singapore	0.0743380	0.0828990‡	0.0790650†‡	0.0752420	0.0781400	0.0543460•	0.0882870†
Spain	0.1436480	0.1625010†	0.1537410†‡	0.1408160	0.1474650	0.0322920•	0.1559140
Sweden	0.1272220	0.1390040†	0.1364530†‡	0.1258640	0.1325020	0.0112540•	0.1379610
Thailand	0.0499110‡	0.0488680	0.0510940†	0.0489480	0.0498050†‡	0.0276390•	0.0465880
Turkey	0.0757640	0.0755440	0.0761250	0.0842020‡	0.0847600†	0.0404750•	0.0830830†‡
UK	0.1242620	0.1339900‡	0.1344120†	0.1242690	0.1245640†‡	0.0003830•	0.1170530
USA	0.0968170	0.1085450‡	0.1093180†	0.0972570	0.0996920	0.0079360•	0.1048940†‡

Notes: †: Best performing model in the row, ‡: Second best performing model in the row,
†‡: Third best performing model in the row, •:Worst performing model in the row.

Note: As can be seen some R^2's are very low. This does not come as a surprise. Andersen and Bollerslev (1998) address the problem of low R^2's, by proving that regression methods will give low R^2 values when daily squared returns measure true volatility, even for optimal GARCH forecasts, because squared returns are noisy estimates of volatility.

3.5 Further categorisation of results (Developed Vs Emerging)

The 25 countries of the sample are from a wide geographical perspective with countries from Europe, Asia, America and Australia. In an attempt to identify any global trends and also evaluate and compare the volatility models in developed and emerging markets the above results are categorised further. First, as before, by the statistical method used to evaluate the forecast performance and then by market classification.[26] In the tables below only the best and worst models are indicated.

In Table 3.8a the ME statistic is reported for all the developed markets of the sample. The results suggest that the Exponential Smoothing model is the best model (in 10 out of 15 countries) with the MA model performing better in four countries and finally the TGARCH in one country. Identifying the worst performing model the HYGARCH is first (in 11 out of 15 counties, although in the case of the USA it is seen as the best model) and then follows the EGARCH model in the rest of the countries. The results do show homogeneity when looking at the best and worst models.

[26] By classification the distinction between developed or emerging markets is implied.

Table 3.8a ME for developed markets

Country/ Model	GARCH	EGARCH	TGARCH	CGARCH	HYGARCH	Ex. Sm	MA
Australia	-0.0000097	-0.0000073	-0.0000107	-0.0000079	0.0000172•	-0.0000117†	-0.0000022
Austria	-0.0000164	-0.0000086	-0.0000113	-0.0000154	-0.0000007•	-0.0000203†	-0.0000069
Belgium	0.0000199	0.0000333•	0.0000186	0.0000185	0.0000127	0.0000112†	0.0000143
Denmark	0.0000181	0.0000216	0.0000176	0.0000177	0.0000292•	0.0000091	0.0000089†
France	0.0000295	0.0000265	0.0000283	0.0000282	0.0000551•	0.0000061†	0.0000153
Germany	0.0000341	0.0000403•	0.0000371	0.0000275	0.0000201	0.0000196	0.0000189†
Hong Kong	-0.0000232	-0.0000341	-0.0000327	-0.0000182	0.0000182•	-0.0001111†	-0.0000145
Ireland	0.0000058	0.0000039	0.0000027	0.0000054	0.0000184•	-0.0000007†	-0.0000007†
Japan	-0.0000077	-0.0000090	-0.0000116	-0.0000070	0.0000318•	-0.0000181†	-0.0000013
Netherlands	0.0000248	0.0000338•	0.0000232	0.0000254	0.0000247	0.0000206	0.0000185†
Singapore	-0.0000132	-0.0000176	-0.0000164	-0.0000073	0.0000015•	-0.0000635†	-0.0000065
Spain	0.0000121	0.0000097	0.0000116	0.0000119	0.0000512•	-0.0000178†	0.0000063
Sweden	0.0000159	0.0000238	0.0000089	0.0000172	0.0000414•	0.0000058†	0.0000111
UK	0.0000094	0.0000122	0.0000014†	0.0000095	0.0000112•	0.0000045	0.0000076
USA	0.0000051	0.0000085•	0.0000055	0.0000054	-0.0000008†	0.0000011	0.0000074

Notes: †: Best performing model in the row, •:Worst performing model in the row.

A similar pattern can be seen when looking at the emerging markets in Table 3.8b. The best by far model is the Exponential Smoothing model with the exception of two countries where the EGARCH performs better. Again the worst performing model is the HYGARCH dominating in all emerging markets except in two where the MA performs poorly. Although the ME statistic is used widely in the literature it does suffer from some technical drawbacks that could affect our conclusions.

Table 3.8b ME for emerging markets

Country/ Model	GARCH	EGARCH	TGARCH	CGARCH	HYGARCH	Ex. Sm	MA
Brazil	-0.0007886	-0.0008597†	-0.0007929	-0.0007535	0.0000175•	-0.0008050	-0.0002151
Chile	-0.0000235	-0.0000259	-0.0000244	-0.0000175	-0.0000105•	-0.0000533†	-0.0000195
India	-0.0000179	-0.0000328†	-0.0000144	-0.0000185	-0.0000004	-0.0000097	0.0000044•
Indonesia	-0.0000370	-0.0000393	-0.0000372	-0.0000101	-0.0000291	-0.0000905†	-0.0000042•
Israel	-0.0000301	-0.0000380	-0.0000395	-0.0000327	0.0000138•	-0.0000431†	-0.0000195
Korea	-0.0000050	-0.0000096	-0.0000181	-0.0000114	-0.0000030•	-0.0001049†	-0.0000067
Malaysia	-0.0000201	-0.0000188	-0.0000225	-0.0000158	-0.0000127•	-0.0001871†	-0.0000586
Philippines	-0.0000244	-0.0000142	-0.0000186	-0.0000235	-0.0000063•	-0.0000670†	-0.0000111
Thailand	-0.0000316	-0.0000404	-0.0000343	-0.0000239	-0.0000031•	-0.0001232†	-0.0000353
Turkey	-0.0000761	-0.0000491	-0.0000790	-0.0000583	0.0001050•	-0.0002102†	-0.0000654

Notes: †: Best performing model in the row, •:Worst performing model in the row.

A more reliable statistic is the MAE, not having the problem of positive and negative values cancelling each other out. The MAE statistic for the developed markets gives a strong HYAGRCH model (in 10 out of 15 countries) and for the remaining countries a strong EGARCH model (in 5 out of 15 countries). For the worst performing model the Exponential Smoothing, unlike before, has the lowest MAE value in all the 15 developed countries.

Table 3.9a MAE for developed markets

Country/ Model	GARCH	EGARCH	TGARCH	CGARCH	HYGARCH	Ex. Sm	MA
Australia	0.0000543	0.0000520	0.0000541	0.0000532	0.0000433†	0.0000583•	0.0000508
Austria	0.0001009	0.0000970	0.0000984	0.0001010	0.0000948†	0.0001078•	0.0000970
Belgium	0.0001375	0.0001328†	0.0001351	0.0001376	0.0001410	0.0001699•	0.0001450
Denmark	0.0001351	0.0001335	0.0001347	0.0001355	0.0001313†	0.0001490•	0.0001412
France	0.0002030	0.0002014	0.0002015	0.0002038	0.0001937†	0.0002439•	0.0002083
Germany	0.0002544	0.0002497†	0.0002508	0.0002544	0.0002583	0.0003031•	0.0002615
Hong Kong	0.0001973	0.0001999	0.0002011	0.0001950	0.0001794†	0.0002532•	0.0001896
Ireland	0.0001141	0.0001137	0.0001150	0.0001148	0.0001080†	0.0001237•	0.0001156
Japan	0.0002088	0.0002060	0.0002089	0.0002089	0.0001960†	0.0002176•	0.0002032
Netherlands	0.0002231	0.0002193†	0.0002210	0.0002229	0.0002236	0.0002779•	0.0002341
Singapore	0.0001446	0.0001440	0.0001461	0.0001421	0.0001381†	0.0001794•	0.0001408
Spain	0.0001760	0.0001743	0.0001746	0.0001770	0.0001631†	0.0002139•	0.0001801
Sweden	0.0002581	0.0002517	0.0002572	0.0002579	0.0002493†	0.0002947•	0.0002593
UK	0.0001317	0.0001297†	0.0001332	0.0001315	0.0001302	0.0001562•	0.0001365
USA	0.0001306	0.0001265†	0.0001275	0.0001306	0.0001319	0.0001464•	0.0001297

Notes: †: Best performing model in the row, •:Worst performing model in the row.

The results are not very different when the emerging market sample is examined, where the HYGARCH model is the best in more than half of the countries. The MA performs well in three countries and the CGARCH model also shows a minimum MAE value. Again, looking at the worst model the Exponential Smoothing dominates with the exception of one country where the EGARCH is the worst. The results in this table are less homogeneous where the first place is shared by more models, although the same conclusions can be drawn.

Table 3.9b MAE for emerging markets							
Country/ Model	**GARCH**	**EGARCH**	**TGARCH**	**CGARCH**	**HYGARCH**	**Ex. Sm**	**MA**
Brazil	0.0008917	0.0009423•	0.0008950	0.0008617	0.0003320†	0.0009149	0.0004463
Chile	0.0000888	0.0000898	0.0000892	0.0000861†	0.0000901	0.0001106•	0.0000876
India	0.0002894	0.0002992	0.0002884	0.0002885	0.0002812†	0.0003260•	0.0002904
Indonesia	0.0002247	0.0002256	0.0002248	0.0002138	0.0002245	0.0002593•	0.0002091†
Israel	0.0001813	0.0001843	0.0001864	0.0001831	0.0001629†	0.0001927•	0.0001765
Korea	0.0003805	0.0003809	0.0003847	0.0003819	0.0003784	0.0004454•	0.0003775†
Malaysia	0.0001027	0.0001023	0.0001048	0.0000991	0.0000964†	0.0002342•	0.0001313
Philippines	0.0002196	0.0002085	0.0002136	0.0002194	0.0002092	0.0002420•	0.0002066†
Thailand	0.0002204	0.0002252	0.0002210	0.0002163	0.0002066†	0.0002823•	0.0002253
Turkey	0.0008415	0.0008288	0.0008423	0.0008248	0.0007568†	0.0009784•	0.0008424
Notes: †: Best performing model in the row, •:Worst performing model in the row.							

Using regression analysis techniques for volatility model evaluation and then categorising the results by market classification the following tables are presented. In Table 3.10a the R^2's of the developed markets are presented. The results are homogenous with very few exceptions. The best model appears to be the EGARCH model (with the exception of two countries where the MA gives the maximum R^2 value). On the other hand the worst performing model in 14 out of 15 countries is the Exponential Smoothing with the exception of Hong Kong where the CGARCH model gives the lowest R^2 value.

Table 3.10a R^2 for developed markets

Country/ Model	GARCH	EGARCH	TGARCH	CGARCH	HYGARCH	Ex. Sm	MA
Australia	0.0439600	0.0661530†	0.0538110	0.0391820	0.0396710	0.0044500•	0.0362370
Austria	0.0692100	0.0753680†	0.0743310	0.0563280	0.0507940	0.0000530•	0.0416130
Belgium	0.1716340	0.2082810†	0.1965110	0.1828970	0.1800390	0.0008600•	0.1149950
Denmark	0.1209240	0.1397390†	0.1296420	0.1180600	0.1277720	0.0000050•	0.0756260
France	0.1936450	0.2230550†	0.2144810	0.1888890	0.1911320	0.0039100•	0.1595550
Germany	0.2079010	0.2360320†	0.2298220	0.2194850	0.2111900	0.0031690•	0.1823330
Hong Kong	0.0508540	0.0710980	0.0538980	0.0481780•	0.0497050	0.0492820	0.0752050†
Ireland	0.0593430	0.0719660†	0.0645930	0.0570530	0.0695980	0.0016650•	0.0490630
Japan	0.0419220	0.0488600†	0.0432740	0.0391380	0.0405330	0.0082910•	0.0422390
Netherlands	0.2322840	0.2538080†	0.2514960	0.2290040	0.2338590	0.0036100•	0.1665290
Singapore	0.0379210	0.0474490	0.0374970	0.0298240	0.0336720	0.0294580•	0.0486200†
Spain	0.1730530	0.1993080†	0.1915030	0.1638660	0.1745510	0.0208830•	0.1466670
Sweden	0.1151830	0.1460580†	0.1404430	0.1139850	0.1156630	0.0027630•	0.1105530
UK	0.1922480	0.2126220†	0.2108590	0.1952620	0.2025500	0.0029040•	0.1404120
USA	0.1257020	0.1807250†	0.1699450	0.1237780	0.1349250	0.0031070•	0.1254140

Notes: †: Best performing model in the row, •:Worst performing model in the row.

In Table 3.10b regression analysis is carried out for the emerging markets of the sample. There is a mixed picture this time when trying to identify the best performing market as several models have maximum R^2 values - the EGARCH model in three countries, the CGARCH, HYGARCH and MA models in two countries respectively and the TGARCH in one country. However, the worst performing model unanimously is the Exponential Smoothing model.

Table 3.10b R^2 for emerging markets

Country/ Model	GARCH	EGARCH	TGARCH	CGARCH	HYGARCH	Ex. Sm	MA
Brazil	0.0397870	0.0506040†	0.040794	0.0355280	0.0236090	0.0135010•	0.0386320
Chile	0.0631800	0.0677110†	0.0644600	0.0627150	0.0078830	0.0075790•	0.0457310
India	0.1167570	0.0969080	0.1132260	0.1320490†	0.1255220	0.0038220•	0.0773180
Indonesia	0.0197210	0.0191530	0.0199410†	0.0131210	0.0155960	0.0015470•	0.0159440
Israel	0.0379870	0.0488220†	0.0394160	0.0409500	0.0401470	0.0147290•	0.0358390
Korea	0.0469650	0.0511460	0.0500490	0.0529290	0.0518960	0.0343120•	0.0710230†
Malaysia	0.0611150	0.0591080	0.0576390	0.0689130	0.0912480†	0.0539510•	0.0630980
Philippines	0.0020900	0.0044520	0.0040200	0.0025430	0.0020700	0.0018680•	0.0315260†
Thailand	0.0835390	0.0775040	0.0934750	0.0982880	0.1021290†	0.0336620•	0.0596080
Turkey	0.1165890	0.1127900	0.1201330	0.1301860†	0.1255700	0.0240740•	0.0921830

Notes: †: Best performing model in the row, •:Worst performing model in the row.

Continuing with the regression analysis the next two tables present the results after running the same regression as above but using logarithms. In Table 3.11a the R^2's of the developed markets show that three models perform better than the rest, so seven countries out of 15 the EGARCH is better, in five countries out of the 15 TGARCH does better and the MA has the maximum R^2 in three countries. There is no dispute as to which models has the worst performance and in all the sample developed countries the Exponential Smoothing has the lowest R^2 values.

Table 3.11a R^2 (log) for developed markets

Country/ Model	GARCH	EGARCH	TGARCH	CGARCH	HYGARCH	Ex. Sm	MA
Australia	0.0455680	0.0685900†	0.0623100	0.0444730	0.0462480	0.0034460•	0.0493230
Austria	0.0398980	0.0407030	0.0420640†	0.0369490	0.0286140	0.0004220•	0.0368900
Belgium	0.1353830	0.1417450	0.1419380†	0.1344420	0.1323460	0.0009860•	0.1268330
Denmark	0.0754470	0.0812940†	0.0754920	0.0784430	0.0802060	0.0003080•	0.0639220
France	0.1083070	0.1283050†	0.1211710	0.1066600	0.1152120	0.0000100•	0.1158010
Germany	0.1685250	0.1806310†	0.1782230	0.1671110	0.1670500	0.0004260•	0.1644680
Hong Kong	0.0695530	0.0796490	0.0709480	0.0729430	0.0752970	0.0807640•	0.0973260†
Ireland	0.0512700	0.0573740†	0.0533990	0.0495920	0.0557620	0.0065380•	0.0542800
Japan	0.0292010	0.0430400	0.0382070	0.0271850	0.0305690	0.0116520•	0.0434610†
Netherlands	0.1663450	0.1710200	0.1730010†	0.1669700	0.1667370	0.0005960•	0.1545450
Singapore	0.0743380	0.0828990	0.0790650	0.0752420	0.0781400	0.0543460•	0.0882870†
Spain	0.1436480	0.1625010†	0.1537410	0.1408160	0.1474650	0.0322920•	0.1559140
Sweden	0.1272220	0.1390040†	0.1364530	0.1258640	0.1325020	0.0112540•	0.1379610
UK	0.1242620	0.1339900	0.1344120†	0.1242690	0.1245640	0.0003830•	0.1170530
USA	0.0968170	0.1085450	0.1093180†	0.0972570	0.0996920	0.0079360•	0.1048940

Notes: †: Best performing model in the row, •:Worst performing model in the row.

The last table of this section Table 3.11b shows the R^2's for the developing markets. Here the picture is a bit more mixed as five models have maximum R^2's. The TGARCH model is the best in three countries, and then follow in two countries the EGARCH, HYGARCH, and MA models and also the GARCH model in one country. The worst model results are also spread across four models, with the Exponential Smoothing having seven minimum values and then one minimum value for the TGARCH, HYGARCH and MA models.

Table 3.22b R^2 (log) for emerging markets

Country/ Model	GARCH	EGARCH	TGARCH	CGARCH	HYGARCH	Ex. Sm	MA
Brazil	0.0218210	0.0250380†	0.0222150	0.0181770	0.0106530•	0.0160050	0.0211340
Chile	0.0294800	0.0320360†	0.0304210	0.0301720	0.0015920	0.0009440•	0.0230740
India	0.0864300†	0.0799710	0.0852160	0.0840990	0.0836900	0.0033960•	0.0717610
Indonesia	0.0138490	0.0098140	0.0140020†	0.0107730	0.0125280	0.0041960	0.0031270•
Israel	0.0367300	0.0383280	0.0397440†	0.0369690	0.0381140	0.0158350•	0.0298770
Korea	0.0600380	0.0624760	0.0648960	0.0612530	0.0625900	0.0453010•	0.0725860†
Malaysia	0.0881460	0.0945220	0.0871060•	0.0960610	0.1041460	0.0935900	0.1043990†
Philippines	0.0151090	0.0140630	0.0156820	0.0148710	0.0161840†	0.0000720•	0.0102600
Thailand	0.0499110	0.0488680	0.0510940†	0.0489480	0.0498050	0.0276390•	0.0465880
Turkey	0.0757640	0.0755440	0.0761250	0.0842020	0.0847600†	0.0404750•	0.0830830

Notes: †: Best performing model in the row, •:Worst performing model in the row.

The tables above suggest developed countries give a less mixed picture when trying to identify the best performing model compared to emerging markets. The overall results suggest that in either developed or emerging markets the asymmetric type GARCH models perform systematically better, compared to the long memory GARCH type models and the more simple models. However, there seems to be little or no dispute as to which model is the worst overall performing model in both developed and emerging markets. It is found that the Exponential Smoothing is the worst performing model in most countries.

Note: As can be observed in several cases very low coefficients of determination are reported. For this reason the significance of the beta coefficient of both the MZ regressions are assessed. The results are presented in appendix 4. In only nine cases and only for the Exponential Smoothing model for the logarithmic MZ regression the beta coefficient was found to be statistically insignificant, and this corresponds to

very low R^2 's. Similarly, for the MZ regression not in logarithmic form in only five cases the beta coefficient for the Exponential Smoothing model was found to be statistically insignificant.

Table 3.12 Summary

Statistical Measure	Models/ Performance	GARCH	EGARCH	TGARCH	CGARCH	HYGARCH	Ex. Sm	MA
ME	Best	0	1	1	0	11	4	9
	Worst	0	7	0	0	7	11	0
MAE	Best	5	0	0	1	16	0	3
	Worst	0	1	0	0	0	24	0
R^2	Best	0	16	1	2	2	0	4
	Worst	0	0	0	1	0	24	0
R^2 **(log)**	Best	1	9	8	0	2	0	5
	Worst	1	0	1	0	1	21	1
Totals	Best	6	26	10	3	31	4	21
	Worst	1	8	1	1	8	80	1

3.6 Conclusion

In an attempt to determine the best and worst performing models the following tables were created. The summary tables give us a snapshot of the exercise's results showing the 'winners' and 'losers' of the exercise comparing seven forecasting volatility models in 25 markets, developed and emerging markets, against four statistical measures of accuracy.

The picture of the best performing model is not a clear one. Ranking the models based on the results the HYGARCH model is first, followed the EGARCH model and the MA, it appears to be one model of each model category, long memory GARCH class

models, asymmetric GARCH class models and the 'simpler' models. Continuing with the ranking of the individual models the TGARCH is fourth, the GARCH fifth the Exponential Smoothing sixth and finally the CGARCH seventh. Ranking the results based on the model classification asymmetric GRACH class models come first, long memory GARCH class models come second and in third place the 'simpler' models. The above ranking however does raise questions especially when extremely low and high results give a mixed overall conclusion. For this reason we categorise the results of Table 3.12 on the basis of the statistical measure used for the forecast evaluation. Tables 3.13 and 3.14 do this. Specifically, when looking at Table 3.13 where the ME and MAE statistical measures of accuracy are used the long memory GARCH class models come first.

Table 3.13 Summary

Statistical measure	Models/ Performance	GARCH	EGARCH	TGARCH	CGARCH	HYGARCH	Ex. Sm	MA
ME	Best	0	1	1	0	11	4	9
	Worst	0	7	0	0	7	11	0
MAE	Best	5	0	0	1	16	0	3
	Worst	0	1	0	0	0	24	0
Totals	Best	6	1	1	1	27	4	11
	Worst	0	8	0	0	7	35	0

Table 3.14 gives the results with respect to the regression analysis techniques employed when the forecasts were evaluated. In this case the first place belongs to the asymmetric class models the EGARCH and TGARCH models.

Table 3.14 Summary

Statistical Measure	Models/ Performance	GARCH	EGARCH	TGARCH	CGARCH	HYGARCH	Ex. Sm	MA
R^2	Best	0	16	1	2	2	0	4
	Worst	0	0	0	1	0	24	0
R^2 **(log)**	Best	1	9	8	0	2	0	5
	Worst	1	0	1	0	1	21	1
Totals	Best	1	25	9	2	4	0	9
	Worst	1	0	1	1	1	45	1

Looking at the losers of the exercise the Exponential Smoothing model comes first, followed by the HYGARCH model, the EGARCH model, the MA and finally in no particular order the GARCH, TAGRCH and CGARCH models. From the losers point of view the Exponential Smoothing is the only one as all the rest of the models come second with the same number of one losing cases.

By performing a further categorization based on the country classification the following conclusions can be drawn. When the least accurate measure of forecast comparisons the ME measure is used for the developed markets the simpler models, more specifically the Exponential Smoothing model in 10 out of 15 countries and the MA in 4 out of 10 countries appear to be the best performers. But once the more accurate measures of forecast comparisons are used the picture changes completely; the Exponential Smoothing model is the worst performing model. The best models appear to be the EGARCH model followed by the HYGARCH model, one asymmetric and one long memory model belonging to the GARCH genre of models. However overall the asymmetric GARCH models (EGARCH and TGARCH) are the best.

When looking at the sample of emerging markets the first observation is that the results are more mixed presenting a less clear picture of the winners and losers of this exercise.

The ME, as in the developed economies, projects the Exponential Smoothing as the winner of the comparisons in 8 out of 10 countries, and the worst model in an equivalent number of countries is the HYGARCH. Excluding the ME from the

analysis a different outcome emerges. The HYGARCH model is the best overall followed by the EGARCH and MA, one asymmetric and one long memory GARCH type model and simple model. In terms of the worst performer the answer is more clear and similar to the one from the developed economies, the Exponential Smoothing is the worst performer.

The volatile nature of emerging markets and their increased sensitivity to local and global events make the modelling and forecasting of volatility a more difficult task. As can be seen above the results in developed economies are more homogeneous and in determining the more accurate models a less challenging job. In emerging markets, due to the variability in the results, drawing conclusions in establishing the best model is not straight forward. On the other hand, homogeneity in identifying the worst performing model for both developed and emerging markets exists, and both samples point to the simple model of Exponential Smoothing gaining this position.

Table 3.15 Summary for developed markets

Statistical measure	Models/ Performance	GARCH	EGARCH	TGARCH	CGARCH	HYGARCH	Ex. Sm	MA
ME	**Best**	0	0	1	0	1	10	4
	Worst	0	4	0	0	11	0	0
MAE	**Best**	0	5	0	0	10	0	0
	Worst	0	0	0	0	0	15	0
R^2	**Best**	0	13	0	0	0	0	2
	Worst	0	0	0	1	0	14	0
R^2 **(log)**	**Best**	0	7	5	0	0	0	3
	Worst	0	0	0	0	0	15	0
Totals	**Best**	0	25	6	0	11	10	9
	Worst	0	4	0	1	11	44	0

Table 3.16 Summary for emerging markets

Statistical measure	Models/ Performance	GARCH	EGARCH	TGARCH	CGARCH	HYGARCH	Ex. Sm	MA
ME	**Best**	0	2	0	0	0	8	0
	Worst	0	0	0	0	8	0	2
MAE	**Best**	0	0	0	1	6	0	3
	Worst	0	1	0	0	0	9	0
R^2	**Best**	0	3	1	2	2	0	2
	Worst	0	0	0	0	0	10	0
R^2 **(log)**	**Best**	1	2	3	0	2	0	2
	Worst	0	0	1	0	1	7	1
Totals	**Best**	1	7	4	3	10	8	7
	Worst	0	1	1	0	9	25	3

4. A backward recursion volatility forecasting exercise[27]

Abstract

A key parameter ignored so far by the academic literature for producing volatility forecasts is the size of the in-sample period used for the forecasts. This chapter aims to answer the question 'how much previous data do we need in order to produce accurate forecasts?' This question introduces the notion of recursive forecasts a topic relatively new within the volatility forecasting literature. The debate is between practitioner/investors and researchers/academics who share different views regarding this question. Respectively, a small in-sample period (small number of observations) is preferred to a large in-sample period (large number of observations) when forecasting volatility. The results show a degree of homogeneity. For most countries of the sample and for the majority of the models large in-sample periods are not necessary for producing the most accurate forecasts supporting the practitioners/investors view; however the models that produce the most accurate forecasts require larger in-sample durations. Furthermore, when taking into account the country classification smaller in-sample durations are required for producing accurate forecasts in emerging markets but more accurate forecasts produced for countries in developed economies.

[27] This chapter was presented as part of a working paper at the BAFA Doctoral Colloquium at Aston Business School on the 11th-12th of April 2011.

4.1 Introduction

Following from the previous chapter there are a number of parameters that affect the accuracy of volatility forecasting and over the years the literature has extensively researched some of these parameters. However, the focus of the literature has mainly been on the type of models used in order to produce accurate volatility forecasts capturing the different features found in the datasets used. Due to this a large number of models have been proposed and many papers have been written comparing those models as can be seen in the previous chapters.

Time series analysis was employed to model volatility forecasting and the debate between the supporters of the 'simpler' models and the GARCH class supporters began and continues even today. A more recent model type introduced was the stochastic volatility models using option prices for their estimation. Another important parameter taken into account by the literature is the data frequency employed. Initially studies used daily, weekly and monthly data but more recent studies look also at intra-daily and high frequency data, again having as the main aim of producing accurate volatility forecasts. A further key parameter for producing volatility forecasts would be the in-sample period used for the forecasts, which will be the main focus of this chapter. The question of *'how much previous data do we need in order to produce accurate forecasts?'* This question introduces the notion of recursive forecasts and more specifically the method of backward recursion which will be used in this chapter. So far the literature has not attempted to answer the question of specifying an ideal in-sample length in order to more accurately produce out-of-sample forecasts.

The tendency in academia in the application of out-of-sample forecasting is to use large in-sample data and often as many observations as possible. On the other hand practitioners tend to use small sample periods not only due to cost and storage restrictions but also because models developed by and used in the finance profession require only a limited amount of data. Two examples of techniques used by practitioners are the first the RiskMetrrics and second the Value at Risk methodologies. The RiskMetrics model was developed by J P Morgan, which when used for forecasting, once the smoothing parameter (λ) is estimated[28] the only variables needed for calculating tomorrow's volatility are today's volatility and today's squared return, both of which are known at the end of trading day. For the Value-at-Risk methodology the minimum period required for backtesting is one year. The purpose of this chapter is to seek an answer to the question above using the some of the models employed in the previous chapter and determining the optimal in-sample length. The rest of this chapter is structured as follows; section 2 presents the data and the methodology, in section 3 the results are presented and analysis is carried out and finally in section 4 some concluding remarks are made.

[28] RM formula: $\sigma^2{}_t = \lambda\sigma_{t-1}^2 + (1-\lambda)\, r_{t-1}^2 = (1-\lambda)\sum_{\tau=1}^{\infty} r_{t-\tau}^2$, λ is estimated to be 0.94 for daily forecasts.

4.2 Data and methodology

4.2.1 Data

The same dataset from chapter 3 is used. All the data are obtained from Datastream market information service. For all the countries daily closing price data from 1 January 1990 to 31 July 2006 are selected and the price indices are converted to returns by the standard method of calculating the log-differences. The descriptive statistics can be found in the data and methodology Chapter 3, section 3.3, table 3.3.

4.2.2 Methodology

For the backward recursion exercise and due to the nature of this exercise different in-sample and out-of-sample periods are used. The main objective of this exercise is to determine the optimal number of in-sample observations to produce the most accurate forecasts, consequently and by definition the in-sample span will vary. From the sample mentioned above the first and last few observations are disregarded to avoid any miscalculations once the 'rolling window' of 60 observations is applied. For the in-sample period the end date will remain fixed on the 29/07/2005 and what will change is the beginning of the in-sample period. The rolling window of 60 observations will then be rolled back to the start of the variable in-sample period producing a forecast for every 60 observations. For each window the Mincer-Zarnowitz (MZ) regression will be run[29]:

[29] The MZ regression is a test regression where inferences regarding individual parameters are less important (the beta coefficient is examined in the point above), with the coefficient of determination being the most important statistic for the purposes of comparison. Due to the potential presence of heteroscedasticity and autocorrelation the Newey-West procedure is used where the standard errors (and t statistics) are corrected for both heteroscedasticity and autocorrelation.

$$\sigma_t^2 = \alpha + \beta h_t^2 + \varepsilon_t \tag{4.1}$$

where the true volatility value is regressed on the forecast value and the coefficient of determination is obtained R^2 for comparison purposes. As a measure of true volatility the procedure of Pagan and Schwert (1990) is followed, where true volatility is used as a proxy by the squared error from the conditional mean model for returns estimated over the whole sample. Each forecast is produced using the models introduced in chapter 3, from the GARCH genre; GARCH(1,1), EGARCH, TGARCH and CGARCH and the representative simple model Moving Average (MA).

Modelling

In this section a small reminder of the models to be used in this chapter are presented. These models have extensively been looked into in literature review section 2.6 (page 42) and in Chapter 3 section 3.3.2 (page 69).

The GARCH(1,1) model of Engle (1982) and Bolerslev (1986) is given by:

$$h_{t+1}^2 = \omega + \alpha \varepsilon_t^2 + \beta h_t^2 \tag{4.2}$$

The Threshold-GARCH (TGARCH) model of Glosten, Jagannathan and Runkle (1993), is given by:

$$h_{t+1}^2 = \omega + \alpha \varepsilon_t^2 + \gamma \varepsilon_t^2 I_t + \beta h_t^2 \tag{4.3}$$

The EGARCH model of Nelson (1991) is given by:

$$\log(h_{t+1}^2) = \omega + \alpha\left|\frac{\varepsilon_t}{h_t}\right| + \gamma\frac{\varepsilon_t}{h_t} + \beta\log(h_t^2) \tag{4.4}$$

The component GARCH (CGARCH) model of Engle and Lee (1999) is specified as:

$$h_{t+1}^2 = q_{t+1} + \alpha(\varepsilon_t^2 - q_t) + \beta(h_t^2 - q_t) \tag{4.5}$$

Moving Average (MA)

Under the moving average method, volatility is forecast by an unweighted average of previously observed volatilities over a particular historical time interval of fixed length:

$$h_{t+1}^2 = \frac{1}{p}\sum_{j=1}^{P}\sigma_j^2 \tag{4.6}$$

where P is the moving average period or 'rolling window'. The choice of this interval is essentially arbitrary, however, the length chosen here equates to a window of 60 days.[30]

To identify the optimal number of in-sample observations the forecast is regressed using Ordinary Least Squares on a constant and a measure of true volatility the squared returns as before (see equation: 4.1). The coefficient of determination (R^2) is obtained for each regression. The maximum value of R^2 signifies the best forecast and

[30] The 60 window is approximately 3 months of trading often regarded as the length of time over which practitioners evaluate their models.

hence indicates the desired in-sample length. The results are presented in both tabular and graphical forms, for all the countries of the sample and for all the models employed. All the calculations are carried out using the statistical software package EViews.

4.3 Results and analysis

The results of the backward recursion forecast exercise are presented below. First a table for each model is presented then followed by a graphical representation of the results. In the first column of each table the sample countries in alphabetical order are found. The next four columns show the length of the in-sample period for which the most accurate forecast is achieved, which was the objective of this exercise. Columns two and three give us respectively the beginning and the end of the in-sample period. As previously mentioned in the methodology section, the end of the in-sample period is fixed at the date of 29/07/2005, whereas the date in column two is variable. Columns four and five illustrate what columns two and three do but in a more comprehensive way. Column four gives the number of observations that where needed in order to obtain the best forecast and column five converts the number of observations into years. The results in the fifth column are derived from the column four after dividing the number of observations by 240, since the data used are daily observations (240 is the number of observations in a year). Finally, in the last column the coefficient of determination R^2 is reported, for comparison purposes.

4.3.1 GARCH(1,1)

Table 4.1 shows the results of the backward recursion exercise using the GARCH model. The range of the observations within the in-sample period is between 179 observations or 0.75 years and 3779 observations, 15.75 years. The smallest in-sample period for producing the best forecast is reported for Israel, where a relatively very small number of observations (179 observations) or only three quarters of a year was used. The reported R^2 value of the regression for Israel is 0.051167. The second smallest in-sample period is recorded for jointly Denmark, India and Malaysia with an in-sample period of 299 observations for producing the accurate forecasts which translates into 1.25 years. The reported R^2's for the three countries respectively are 0.054379, 0.244684 and 0.033788. It is worth mentioning that so far all the countries are from developing economies. In third joint place, in terms of smallest in-sample period, with 359 observations or a year and half are Hong Kong, Indonesia, Sweden and the UK with R^2's of the values of 0.021559, 0.049918, 0.184684 and 0.145232. On the other hand the three longest in-sample periods for producing the best forecast are reported for Brazil, Turkey and the USA with respectively an in-sample period of 3779 observations or 15.75 years, 3239 observations or 13.50 years and 2699 observations or 11.25 years. In the same order the R^2's are 0.050951, 0.040961 and 0.023786.

The average in-sample period for all the 25 countries of the sample when using the GARCH model is 1108 observations or in terms of years 4.6.

Table 4.1 Results of backward recursion exercise for GARCH

Country	In-sample Start Date	In-sample End date	In-sample size Number of observations	In-sample size Years	R^2
Australia	05/08/2002	29/07/2005	779	3.25	0.072
Austria	22/12/2003	29/07/2005	419	1.75	0.092
Belgium	14/04/2003	29/07/2005	599	2.50	0.135
Brazil	04/02/1991	29/07/2005	3779	15.75	0.050
Chile	07/07/2003	29/07/2005	539	2.25	0.024
Denmark	07/06/2004	29/07/2005	299	1.25	0.054
France	29/09/2003	29/07/2005	479	2.00	0.099
Germany	04/11/1996	29/07/2005	2279	9.50	0.076
Hong Kong	15/03/2004	29/07/2005	359	1.50	0.021
India	07/06/2004	29/07/2005	299	1.25	0.244
Indonesia	15/03/2004	29/07/2005	359	1.50	0.049
Ireland	04/11/1996	29/07/2005	2279	9.50	0.068
Israel	22/11/2004	29/07/2005	179	0.75	0.051
Japan	07/07/2003	29/07/2005	539	2.25	0.048
Korea	26/11/2001	29/07/2005	959	4.00	0.024
Malaysia	07/06/2004	29/07/2005	299	1.25	0.033
Netherlands	19/06/1995	29/07/2005	2639	11.00	0.092
Philippines	11/09/1995	29/07/2005	2579	10.75	0.092
Singapore	22/12/2003	29/07/2005	419	1.75	0.093
Spain	22/12/2003	29/07/2005	419	1.75	0.116
Sweden	15/03/2004	29/07/2005	359	1.50	0.184
Thailand	07/07/2003	29/07/2005	539	2.25	0.029
Turkey	01/03/1993	29/07/2005	3239	13.50	0.040
UK	15/03/2004	29/07/2005	359	1.50	0.145
USA	27/03/1995	29/07/2005	2699	11.25	0.023

In the graphs below the R^2's are plotted to graphically depict the results of the previous table. Some graphs appear to be flatter than other and some have distinct peaks. The highest peak on each graph indicates where the best forecast is, against the horizontal axis indicating the length of the in-sample period. The closer to the right or the end of the line the peak is positioned on each graph the smaller the in-sample period is required for producing accurate forecasts. In some of the graphs the highest peak is easy to identify however where the line is flatter this is not the case. It is worth noting that flatter lines do not imply less volatility because what is being looked into and what is represented by the line is the accuracy of the volatility forecasts.

Looking at the graphs a pattern seems to emerge, the flatter lines are observed for the developed economies whereas peaked lines are more visible for the developing markets. So far and while the GARCH (1,1) is used in the forecast exercise, on average smaller in-sample periods for accurate forecasts are observed in developed markets. The highest R^2's values are in the range of 25% and 2%, for India and Hong Kong, one emerging and one developed Asian economy respectively. On average the R^2 value for the developed economies is higher than that for the developing economies.

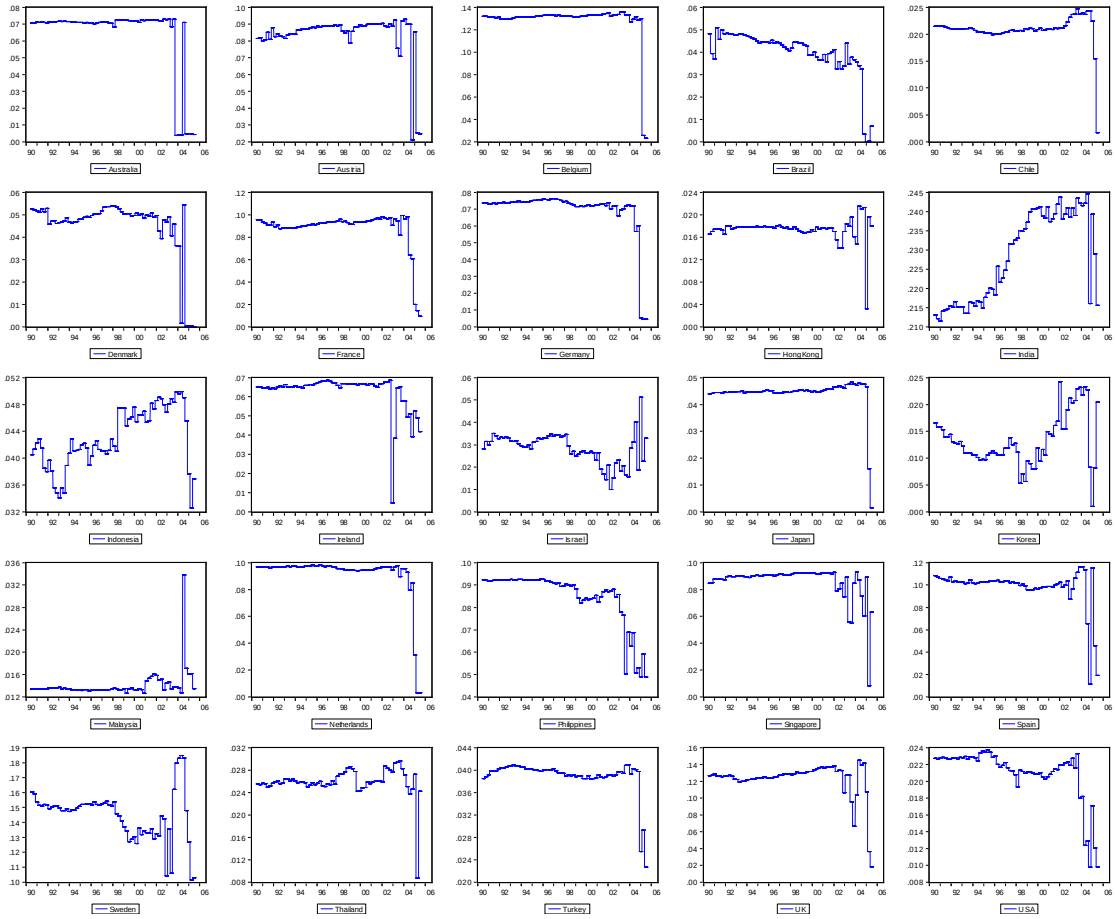
Some developed economies are worth being looked into in more detail as unusually long in-sample periods are reported. Specifically the cases of Germany, Ireland, Netherlands and the USA. The graphs for Germany Ireland and the Netherlands appear to be relatively flat with the exception of the very low minimum point in 2003 for Ireland[31], making it difficult from the graph to find the highest value for the R^2. Therefore, although long in-sample periods are reported, due to the relatively flat line there are minor differences in the value of the R^2. In the case of the USA the R^2 plot has two distinct peaks one around 1995 and the other in 2003.

Overall and as discussed so far there is no trade off between the in-sample length and the accuracy of the forecast when the GARCH (1,1) model is used. On average shorter in-sample periods are observed for developed markets and more accurate volatility forecasts are also observed for the developed economies. Mixed results are reported when using GARCH (1,1) in an attempt to answer the question of how long in-sample periods should be used in order to obtain accurate forecasts. The tendency

[31] This could be treated as an outlier.

in academia is to use large data sets in contrast to practitioners who due to cost minimisation use smaller data sets. On average when deciding on the length of an in-sample period the first criterion is to determine whether the sample country used is from a developed or developing economy. Less observations (by a year) are required when using daily data for developed economies and the specified model for volatility estimation is the GARCH (1,1) model. In the sections that follow further comparisons will be able to be made in terms of model used.

Graph 4.1
Graphical representation of results of backward recursion exercise with GARCH



4.3.2 EGARCH

Next the backward recursion exercise is performed using the second generation of GARCH models the EGARCH and TGARCH models. The analysis will begin by looking into the EGARCH calculations. Table 4.2 shows the results in the same format as before. The range of the in-sample periods start from 119 daily observations, half a year with a maximum of 2699 observations, 11.25 years. The smallest in-sample period for producing the best forecast is seen to be for Korea with 119 daily observations -half a year. The R^2 of the regression for the forecast performance analysis for Korea is 0.08182. The second smallest in-sample period is reported for Japan with 239 observations representing one year, with a relatively high R^2 of 0.121862. In third joint place in terms of the smallest in-sample period for producing accurate forecasts come Austria, Brazil, India, Netherlands, Sweden and Turkey with an in-sample span of 299 observations, a year and a quarter. In the same order the R^2's are 0.155783, 0.040239, 0.326213, 0.149255, 0.330333 and 0.084662. Looking at the longest in-sample periods Hong Kong dominates with a relative to the rest of the countries very high number of observations 2699 or 11.25 years, and this is put into perspective when then next longest in-sample period with almost half of the number of observations is for Australia with 1319 daily observations, five and a half years, followed closely by Ireland with 1197 daily observations or 4.99 years. The R^2's for the above three countries are in the same order 0.02660, 0.093368 and 0.116904.

The average number of observations for all the 25 countries of the sample when using the EGARCH model for the backward recursion exercise is 635 observations or 2.65 years. This is almost the half the amount of data needed to produce accurate forecasts

when using the GARCH model. It is worth noting that if the extreme result for Hong Kong is regarded as an outlier the average in-sample period reduces to 549 observations or 2.29 years.

Table 4.2 Results of backward recursion exercise for EGARCH

Country	In-sample Start Date	In-sample End date	In-sample size Number of observations	In-sample size Years	R^2
Australia	10/07/2000	29/07/2005	1319	5.50	0.093
Austria	15/03/2004	29/07/2005	359	1.50	0.126
Belgium	07/06/2004	29/07/2005	299	1.25	0.155
Brazil	26/11/2001	29/07/2005	959	4.00	0.082
Chile	07/06/2004	29/07/2005	299	1.25	0.040
Denmark	22/12/2003	29/07/2005	419	1.75	0.066
France	22/12/2003	29/07/2005	419	1.75	0.168
Germany	15/03/2004	29/07/2005	359	1.50	0.147
Hong Kong	27/03/1995	29/07/2005	2699	11.25	0.026
India	07/06/2004	29/07/2005	299	1.25	0.326
Indonesia	20/01/2003	29/07/2005	659	2.75	0.077
Ireland	25/12/2000	29/07/2005	1197	4.99	0.116
Israel	22/12/2003	29/07/2005	419	1.75	0.091
Japan	30/08/2004	29/07/2005	239	1.00	0.121
Korea	14/02/2005	29/07/2005	119	0.50	0.081
Malaysia	26/11/2001	29/07/2005	959	4.00	0.019
Netherlands	07/06/2004	29/07/2005	299	1.25	0.149
Philippines	18/02/2002	29/07/2005	899	3.75	0.085
Singapore	26/11/2001	29/07/2005	959	4.00	0.105
Spain	07/07/2003	29/07/2005	539	2.25	0.148
Sweden	07/06/2004	29/07/2005	299	1.25	0.330
Thailand	29/09/2003	29/07/2005	479	2.00	0.035
Turkey	07/06/2004	29/07/2005	299	1.25	0.084
UK	29/09/2003	29/07/2005	479	2.00	0.223
USA	14/04/2003	29/07/2005	599	2.50	0.090

There are obvious differences in the graphical representation of the results when the EGARCH model is used in the backward recursion exercise compared to the previous results obtained with the use of the GARCH (1,1) model. The first main difference is that the graphs appear less flat with very high and very low peaks present and in almost all the graphs. The highest peaks representing the R^2 are towards the end of the line, on the right hand side of the graph, confirming that when the EGARCH model is used for forecasting purposes smaller in-sample durations -smaller number

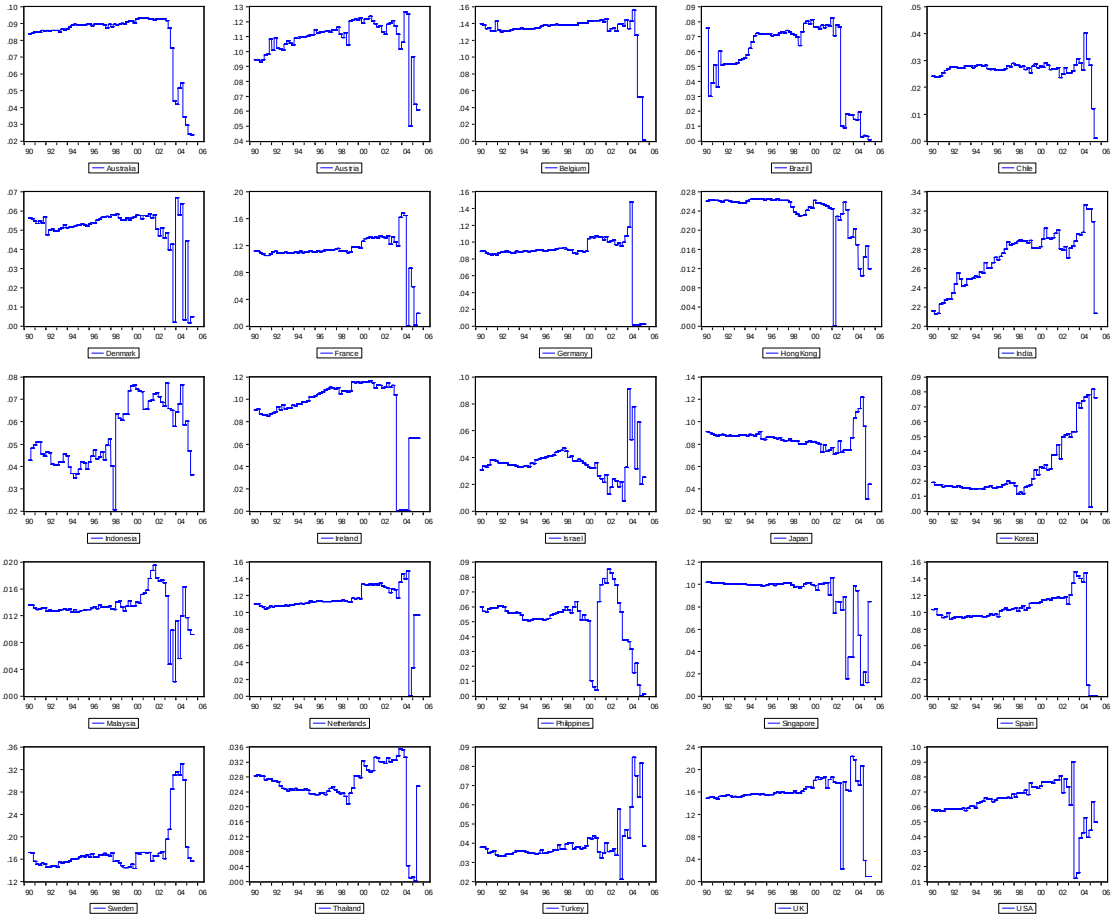
observations, are required in order to produce accurate forecasts. Due to the evident lack of flatness of the graphs, it is easy to determine graphically the highest value of the R^2's.

The case of Hong Kong is a peculiar one compared to the rest of the sample countries and as can be seen from the graphs, the higher R^2 is at the beginning of the in-sample period although there are some high peaks towards the end -it is also worth mentioning the very low R^2 values appear between 2001 and 2002. This case can be considered an extreme one and if excluded as an outlier this improves significantly the overall performance of the EGARCH model in requiring small in-sample periods for producing accurate volatility forecasts.

From the graphs there are no significant differences between emerging markets and developed economies apart from the initial relative flatness of the R^2 plot in some of the graphs. It is not clear cut this time from the graphs, as was when the GARCH (1,1) model was used, which graphs are from developed and which from developing economies. Making comparisons between the developed and emerging markets there is a distinction in terms of average R^2's and in-sample duration with the emerging countries performing better in terms of accuracy with an average R^2 value of 14% and the emerging economies with an average of 9%. In addition when looking at the in-sample durations, less data is required for producing accurate forecasts in the sample of the developing countries, 2.25 years, compared to 3 years of daily data required from the sample of the developed economies. This outcome comes as a surprise and is contrary to the previous findings.

The overall average R^2 value for all countries is between 33% and 2% -for Sweden and Malaysia respectively and with an average value of 12% which is the highest average reported for the purpose of this exercise making the EGARCH model the most accurate in this exercise.

Graph 4.2
Graphical representation of results of backward recursion exercise with EGARCH



4.3.3 TGARCH

Next the backward recursion exercise is performed using the TGARCH model. Table 4.3 presents the results in a similar way as before. The ranges of the in-sample periods

are from 59 observations to 2879, or in years from a quarter of a year up to 12 years. The smallest in-sample period for producing the best forecast is for Turkey with 59 daily observations, 0.25 of a year, and with a reported R^2 of 0.067369. Then follow India and Korea with an in-sample period of 179 observations or 0.75 of a year. The respective R^2's are 0.302553 (an unusually relative high value) and 0.080256. The third smallest in-sample period is for Japan with a period of 239 observations or a year and with a coefficient of determination of 0.102163. On the other hand the longest in-sample periods are for Hong Kong with 2879 observations, or 12 years and with a R^2 of 0.02416, Ireland with 1199 observations or 5 years and with a reported R^2 of 0.104805 and Singapore with 998 observations, 4.16 years and a R^2 of 0.0104681.

While using the TGARCH model in the backward recursion technique the average number of observations for the whole sample required for producing an accurate forecast is 541 or 2.26 years. Compared to the previously used models namely the GARCH and EGARCH, the TGARCH model requires a smaller in-sample period to produce a more accurate forecast. The average number of observations when using the GARCH model is 1108 and when using EGARCH it is 635 observations.

Table 4.3 Results of backward recursion exercise for TGARCH

Country	In-sample Start Date	In-sample End date	In-sample size Number of observations	In-sample size Years	R^2
Australia	07/07/2003	29/07/2005	539	2.25	0.092
Austria	07/06/2004	29/07/2005	299	1.25	0.114
Belgium	07/06/2004	29/07/2005	299	1.25	0.164
Brazil	20/01/2003	29/07/2005	659	2.75	0.073
Chile	07/06/2004	29/07/2005	299	1.25	0.033
Denmark	07/06/2004	29/07/2005	299	1.25	0.074
France	15/03/2004	29/07/2005	359	1.50	0.172
Germany	15/03/2004	29/07/2005	359	1.50	0.132
Hong Kong	18/07/1994	29/07/2005	2879	12.00	0.024
India	22/11/2004	29/07/2005	179	0.75	0.302
Indonesia	20/01/2003	29/07/2005	659	2.75	0.062
Ireland	25/12/2000	29/07/2005	1199	5.00	0.104
Israel	07/06/2004	29/07/2005	299	1.25	0.079
Japan	30/08/2004	29/07/2005	239	1.00	0.102
Korea	22/11/2004	29/07/2005	179	0.75	0.080
Malaysia	30/08/2004	29/07/2005	239	1.00	0.035
Netherlands	07/06/2004	29/07/2005	299	1.25	0.143
Philippines	28/10/2002	29/07/2005	719	3.00	0.100
Singapore	02/10/2000	29/07/2005	998	4.16	0.104
Spain	15/03/2004	29/07/2005	359	1.50	0.159
Sweden	29/09/2003	29/07/2005	479	2.00	0.319
Thailand	07/06/2004	29/07/2005	299	1.25	0.050
Turkey	09/05/2005	29/07/2005	59	0.25	0.067
UK	22/12/2003	29/07/2005	419	1.75	0.204
USA	18/02/2002	29/07/2005	899	3.75	0.064

In most cases, looking at the graphs, the highest peak can be detected, however it is worth mentioning that a relative flatness of the R^2 line can be seen for the sample of developed economies, but not as flat in the case when the GARCH (1,1) model was used. To some extent a distinction can be made between the developed and developing markets in the sample by looking at the graphs, flatter lines are observed in developed economies and less flat for developing. As a reminder, a less flat line does not imply more volatility.

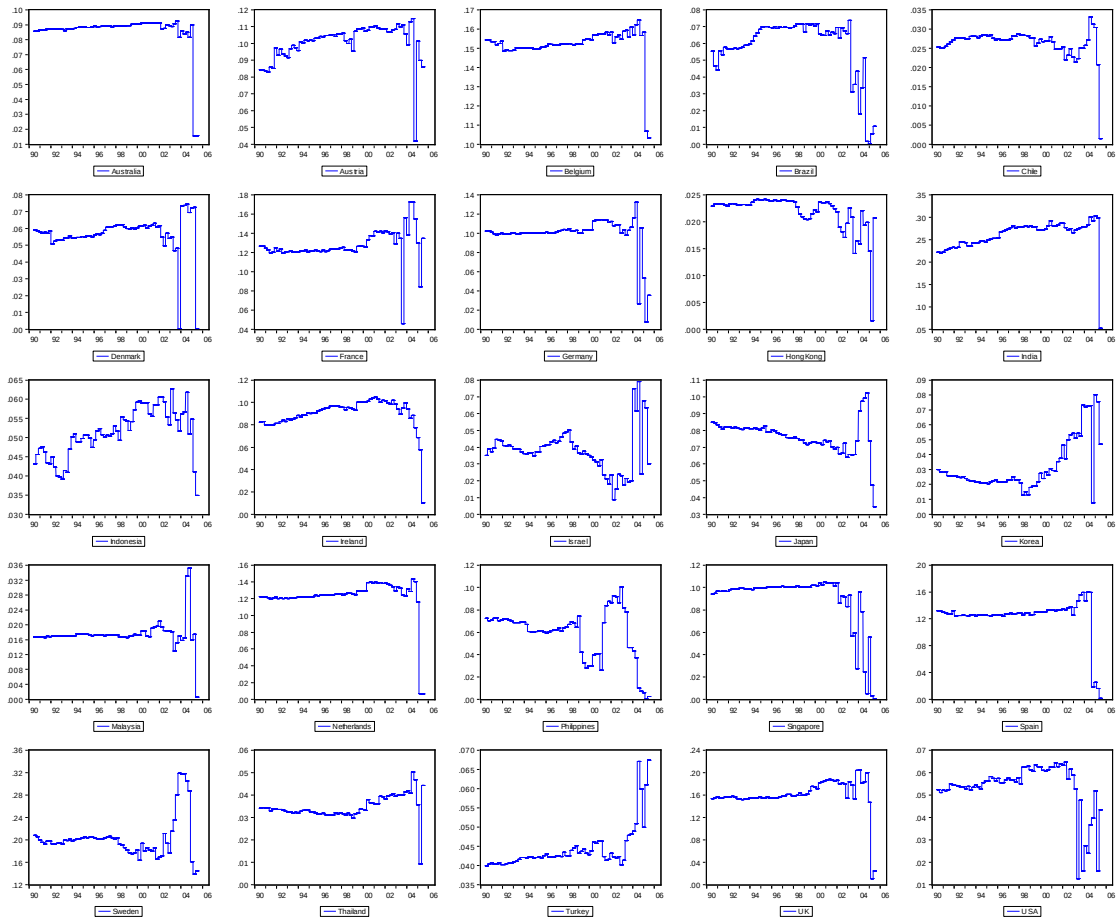
Once more the case of Hong Kong appears to be a peculiar one. Not only does the reported in-sample period appear to be significantly long and falls out of line

compared with the other in-sample periods for the estimation of accurate forecasts, but there are several high peaks on the curve.

Comparing the findings from the graphs of the emerging and developed economies, similarly when the EGARCH model was used for estimation, the average R^2 for the developed countries is 13% but for the developing economies it is approximately 9%. The range for the R^2 is 32% and 2.4% for Sweden and Hong Kong respectively. The same pattern with respect to the in-sample durations also appears. Longer on average in-sample periods are reported for developed economies compared to the average of the developing economies, even if the extreme value of Hong Kong is excluded when averages are estimated. Asymmetric GARCH models seem to require smaller in-sample periods for accurate volatility forecasts however in terms of accuracy the developed economies the asymmetric models do a better job in the developed economies. Specifically on average 2.75 years of daily data are needed in order to produce accurate forecasts in developed markets and 1.5 years of daily data are required when the sample consists of emerging markets.

In terms of model accuracy the asymmetric GARCH models perform better than the GARCH (1,1) model. Between EGARCH and TGARCH models more accurate forecasts are obtained with the use of EGARCH. On the other hand, shorter in-sample periods for accurate volatility estimation in ascending order are needed for TGARCH, EGARCH and finally GARCH.

Graph 4.3
Graphical representation of results of backward recursion exercise with TGARCH



4.3.4 CGARCH

The last model from the ARCH genre of models, capturing the long memory element of the data used in the backward recursion exercise is the CGARCH. Table 4.4 below shows the results of the exercise followed by a graphical presentation of the results. The ranges of the in-sample period for producing accurate forecasts start from 59 daily observations or 0.25 of a year, to 3960 daily observations or 16.5 years. Hong Kong produced the smallest in-sample period of 59 observations with a coefficient of

determination of 0.038542. Israel gives us an in-sample period of 179 observations or three quarters of a year and a R^2 of 0.082662. Finally the third smallest in-sample period is given for Malaysia with 239 observations also representing a year. The R^2 for Malaysia is 0.035851. The countries that needed the largest in-sample periods for producing accurate forecasts are Sweden, Germany and Spain requiring in-sample periods of respectively 3960 observations or 16.5 years, 3899 observations or 16.25 years, and 3839 or 16 years. The R^2 for Sweden is 0.113074, for Germany 0.084155 and for Spain 0.170439.

The average number of observations required for obtaining accurate forecasts in our sample is 1449 or 6 years. This is higher than all the previously used models.

Table 4.4 Results of backward recursion exercise for CGARCH

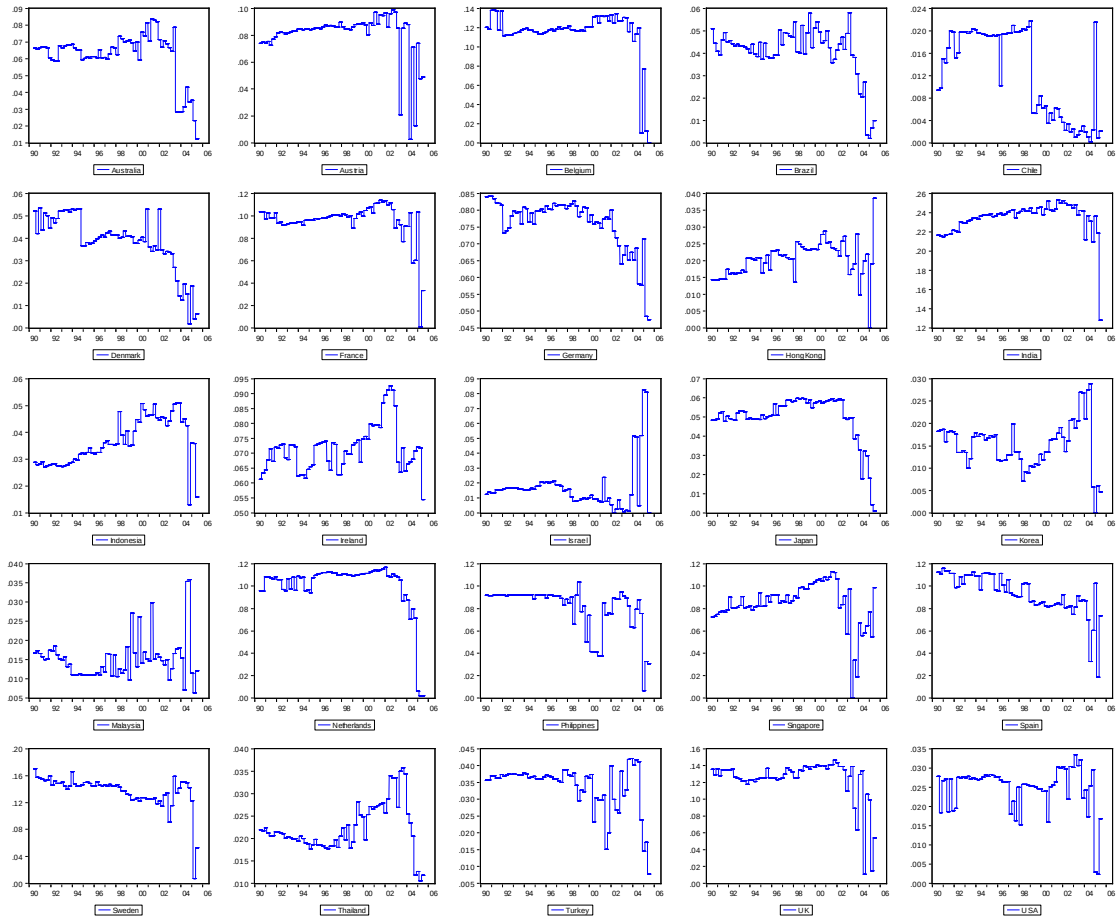
Country	In-sample Start Date	In-sample End date	In-sample size Number of observations	In-sample size Years	R^2
Australia	19/03/2001	29/07/2005	1139	4.75	0.083
Austria	05/08/2002	29/07/2005	779	3.25	0.098
Belgium	04/02/1991	29/07/2005	3779	15.75	0.138
Brazil	20/01/2003	29/07/2005	659	2.75	0.058
Chile	30/11/1998	29/07/2005	1739	7.25	0.021
Denmark	12/11/1990	29/07/2005	3839	16.00	0.053
France	11/06/2001	29/07/2005	1079	4.50	0.114
Germany	20/08/1990	29/07/2005	3899	16.25	0.084
Hong Kong	09/05/2005	29/07/2005	59	0.25	0.038
India	11/06/2001	29/07/2005	1079	4.50	0.253
Indonesia	07/07/2003	29/07/2005	539	2.25	0.051
Ireland	13/05/2002	29/07/2005	839	3.50	0.092
Israel	22/11/2004	29/07/2005	179	0.75	0.082
Japan	07/09/1998	29/07/2005	1799	7.50	0.059
Korea	07/06/2004	29/07/2005	299	1.25	0.028
Malaysia	30/08/2004	29/07/2005	239	1.00	0.035
Netherlands	26/11/2001	29/07/2005	959	4.00	0.116
Philippines	30/11/1998	29/07/2005	1739	7.25	0.103
Singapore	11/06/2001	29/07/2005	1079	4.50	0.113
Spain	12/11/1990	29/07/2005	3839	16.00	0.116
Sweden	25/05/1990	29/07/2005	3960	16.50	0.170
Thailand	07/07/2003	29/07/2005	539	2.25	0.035
Turkey	29/09/2003	29/07/2005	479	2.00	0.042
UK	03/09/2001	29/07/2005	1019	4.25	0.146
USA	20/01/2003	29/07/2005	659	2.75	0.033

There appears to be little to non flatness in the graphs when the CGARCH model is used. Consequently, it is not always obvious where the maximum R^2 is situated? What is also not clear from the graphs is whether the market is developed or developing, as there is no distinct pattern.

The range of the R^2 values is between 25% for India and 2% for Chile. However, the average R^2's are low compared to the asymmetric GARCH models but not to the GARCH (1,1) model. The mean R^2 value for the developed economies is 9.7% and for the emerging markets 7.1 %, indicating better accuracy when the sample data is from the developed economies.

The average in-sample period for producing accurate forecasts for the developed countries is 7.9 years in daily data compared to the average 3.1 years of daily data needed for accurate forecasts when the sample is of emerging markets. This same pattern was detected when the EGARCH and TGARCH models were used but not when the GARCH (1,1) was used. It must be noted that such a big difference in the in-sample periods is reported for the first time.

Graph 4.4
Graphical representation of results of backward recursion exercise with CGARCH



4.3.5 Moving Average

As a representative model from the more 'simple' models the Moving Average is used. In table 4.5 we see the results of the backward recursion exercise. The ranges of the in-sample periods are from 59 observations a quarter of a year, for which 12 out of the 25 countries of the sample give us this minimum in-sample value, however the maximum in-sample period is of 3779 observations which is 15.75 years. The smallest in-sample period for producing an accurate forecast is reported for Chile, France, Germany, Hong Kong, India, Israel, Korea, Netherlands, Philippines,

Singapore, UK and the USA. The respective coefficients of determination, R^2's are: 0.015636, 0.068092, 0.63402, 0.036438, 0.136312, 0.031622, 0.003966, 0.082602, 0.067123, 0.072866, 0.079228 and 0.040962. The second smallest in-sample period is reported for Sweden with 119 observations, half a year and with a R^2 of 0.102141. The third smallest in-sample period for producing accurate forecasts is 179 observations, three quarters of a year, with jointly three countries Austria, Brazil and Spain with their respective R^2's of 0.047151, 0.02996 and 0.070689. The largest in-sample periods are reported first for Indonesia[32] with 3779 observations, 15.75 years and a R^2 of 0.012249. The second largest in-sample period is less than half of that of Indonesia with 1619 observations, 6.75 years for Thailand and a R^2 of 0.048456. The third largest in-sample period is 723 observations, three years, for Australia and with a R^2 of 0.062514.

The average number of observations required for producing accurate forecasts using the MA model in the backward recursion exercise is 388 observations or 1.62 years. If however we take out the two extreme values reported above for Indonesia and Thailand the average becomes 187.2 observations or 0.78 years.

[32] Could be treated as an outlier.

Table 4.5 Results of backward recursion exercise for MA

Country	In-sample Start Date	In-sample End date	In-sample size Number of observations	In-sample size Years	R^2
Australia	28/10/2002	29/07/2005	723	3.01	0.062
Austria	22/11/2004	29/07/2005	179	0.75	0.047
Belgium	15/03/2004	29/07/2005	359	1.50	0.103
Brazil	22/11/2004	29/07/2005	179	0.75	0.029
Chile	09/05/2005	29/07/2005	59	0.25	0.015
Denmark	29/09/2003	29/07/2005	483	2.01	0.038
France	09/05/2005	29/07/2005	59	0.25	0.068
Germany	09/05/2005	29/07/2005	59	0.25	0.063
Hong Kong	09/05/2005	29/07/2005	59	0.25	0.036
India	09/05/2005	29/07/2005	59	0.25	0.136
Indonesia	04/02/1991	29/07/2005	3779	15.75	0.012
Ireland	07/06/2004	29/07/2005	299	1.25	0.073
Israel	09/05/2005	29/07/2005	59	0.25	0.031
Japan	07/06/2004	29/07/2005	299	1.25	0.045
Korea	09/05/2005	29/07/2005	59	0.25	0.003
Malaysia	15/03/2004	29/07/2005	359	1.50	0.025
Netherlands	09/05/2005	29/07/2005	59	0.25	0.082
Philippines	09/05/2005	29/07/2005	59	0.25	0.067
Singapore	09/05/2005	29/07/2005	59	0.25	0.072
Spain	22/11/2004	29/07/2005	179	0.75	0.070
Sweden	14/02/2005	29/07/2005	119	0.50	0.102
Thailand	17/05/1999	29/07/2005	1619	6.75	0.048
Turkey	22/12/2003	29/07/2005	419	1.75	0.058
UK	09/05/2005	29/07/2005	59	0.25	0.079
USA	09/05/2005	29/07/2005	59	0.25	0.040

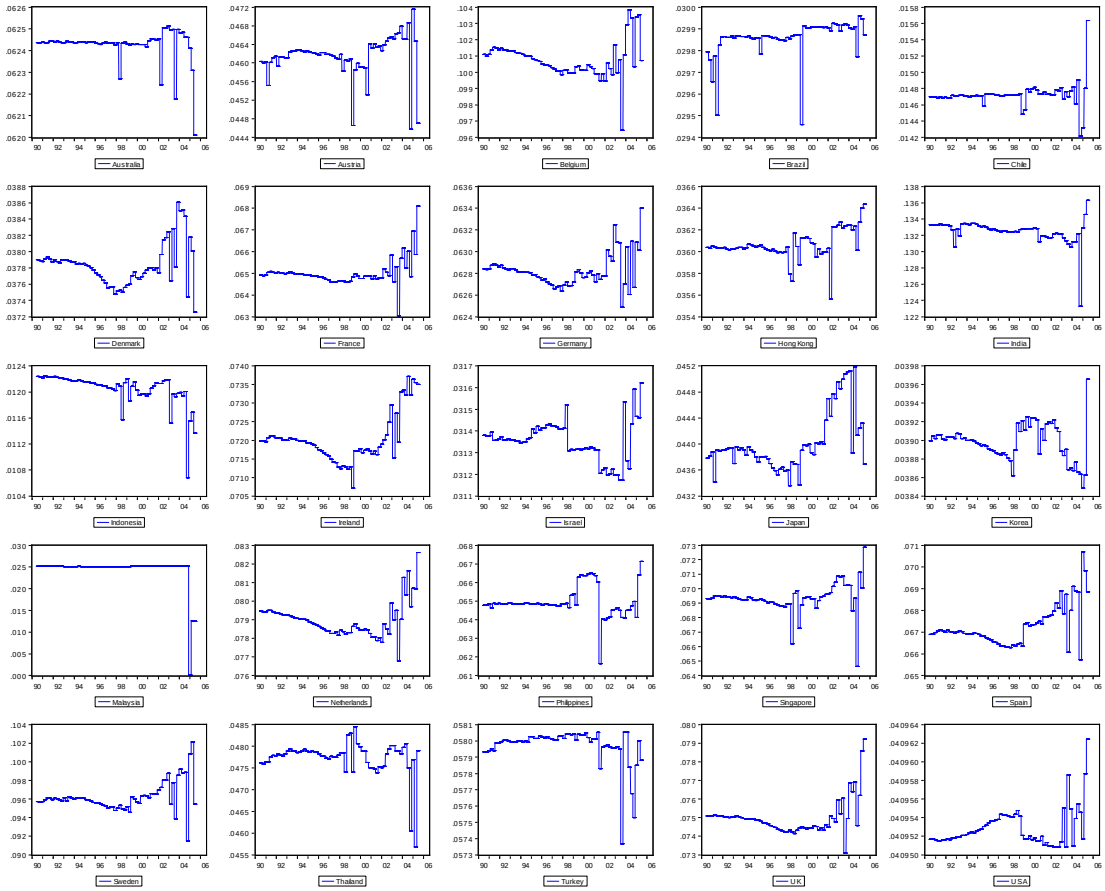
Identifying graphically the highest R^2 value is a difficult task as the MA model produces graphs with several peaks and little flatness. It is worth mentioning the case of Malaysia however where there is just a straight line representing the R^2 plot, but this is just one exception. There also appears to be no distinct pattern for identifying the developed economies from the emerging markets.

In terms of accuracy the simple MA model has done rather poorly compared to the more advanced GARCH type models. For example the maximum R^2 value is 13% reported for India. Interestingly this is the third time out five (five models are sued in this exercise) where India gives the highest R^2 value, high coefficients of

determination were reported when the GARCH and CGARCH models were applied, and Sweden for the remaining two models (EGARCH and TGARCH). The average R^2 value for the emerging economies is 6.5% and for the developing slightly less accurate 4.3%.

When looking at the in-sample durations it is evident that the MA model does not require long datasets for producing accurate forecasts. The mean number of daily observations as mentioned above is 388 or in years 1.6. Breaking down this figure the developed economies require a mean 0.85 of a year in daily observations and the developing economies 2.77 years of daily data. A smaller in-sample period is required for the developed markets when calculating volatility forecasts which are also more accurate compared to when data from emerging markets are used, concluding that there is no trade off between accuracy of volatility estimation and longer datasets required for estimation as was also the case of the GARCH (1,1) model.

Graph 4.5
Graphical representation of results of backward recursion exercise with MA



4.4 Conclusion

The purpose of this exercise was to seek an answer to the question of identifying the appropriate length for the in-sample period when forecasting volatility. Simply put '*how much previous data should be used in order to produce accurate volatility forecasts*'? It is the first time within the finance literature that this question is raised for which mainly two deferring views exist. In academia the tendency is to use large datasets, as much data as possible in order to estimate a forecast model in contrast to

finance practitioners who tend to use small datasets mainly due to costing issues and data storage.

In attempting to answer the above question the models used in the previous chapter are employed and a backward recursive technique is used. The end date of the in-sample period is fixed (in this exercise on the date 29/07/2005) and then using a rolling window of 60 observations which is rolled back to the start of the variable in-sample period producing a forecast for every 60 observations (one quarter) for a selection of models used. More specifically the GARCH (1,1) the first generation of the ARCH genre, two asymmetric GARCH models the EGARCH and TGARCH, a representative GARCH model for capturing the long memory affect the CGARCH and a representative simple model the Moving Average (MA). To identify the optimal number of in-sample observations the Mincer-Zarnowitz (1969) methodology is followed where the forecast estimate is regressed using Ordinary Least Squares on a constant and on a measure of true volatility for the coefficient of determination to be obtained for each regression. The maximum value of R^2 signifies the best forecast and hence indicates the desired in-sample length.

The shorter in-sample periods for producing accurate forecasts were reported for the MA model with an average 388 daily observations (1.62 years) followed by both the asymmetric GARCH models in the order of the TGARCH and then the EGARCH model with respectively an average of 541 and 635 daily observation (2.26 and 2.65 years), followed by the GARCH (1,1) with an average of 1108 daily observations (4.6 years) and finally by the long memory GARCH model the CGARCH with an average of 1449 daily observation (6 years).

In terms of accuracy of forecast estimation the MA, with shortest in-sample period gives the lowest average R^2 value. The lowest R^2 value is given by the GARCH (1,1) model, followed by the CGARCH. Then follow the TGARCH and EGARCH models with the highest and hence more accurate volatility estimations.[33]

Further analysis was carried throughout the chapter to identify possible trends between market/economy classification and performance with respect to in-sample duration and accuracy of forecasts.[34] The in-sample duration required for producing accurate forecasts is smaller when investigating emerging markets and the preferred models are the asymmetric GARCH models (EGARCH or TGARCH)[35] and then the long memory GARCH model namely the CGARCH model. In the cases where the GARCH (1,1) models and the MA are used smaller in-sample periods are required for the estimation of accurate forecasts. However, when looking at the accuracy of the forecasts produced in all cases a better job is done when the data is from developed economies.

[33] Note: Although the precise point for determining the optimal in-sample length –the maximum value for the coefficient of determination, might be short or long for different sample countries, often there are very small differences between the values of the coefficient of determination. This becomes more obvious when looking at the graphs. Therefore, larger in-sample periods can often be preferred. This issue, could be due to structural breaks (the sample period is over 15 years of daily data), which would support smaller optimal periods. From an economic perspective long memory breaks caused by market frictions or trader/investor behaviour –this could be the topic for further research.

[34] The distinction between emerging and developed was based on the FTSE classification. Other potentially interesting factors that could be further explored for example market liquidity, regulatory framework etc.

[35] Both EGARCH and TGARCH capture asymmetries which are known to be important (Black, 1976; Nelson, 1991; Glosten et al., 1993; Bolerslev et al. 1994). In equity returns, which is what we have in this case, such asymmetries are attributed to leverage effects or the alternative view, the volatility feedback hypothesis Brooks (2008). The Exponential GARCH (EGARCH) model due to its logarithmic form does not impose coefficient constraints and thus estimation is improved. On the other hand the Threshold GARCH (TGARCH) model captures the asymmetry effect with a dummy variable which allows for the leverage effect.

Overall shorter in-sample periods of up to 3 years are deemed enough for producing accurate volatility forecasts for three out of the five selected models of this exercise (MA, TGARCH and EGARCH). Longer in-sample periods by almost 2 years are required for the GARCH model and further 2 years are required for the CGARCH model. This finding appears to be more in line with the view of the practitioners; however there appears to be a trade off between the in-sample length and accuracy of forecasts. The models requiring longer in-sample durations produce more accurate forecasts.

5. A Value-at-Risk volatility forecasting exercise[36]

Abstract

The RiskMetrics model is preferred within the investment community due to it its simplicity over the more academic GARCH models. Although academic research has shown the superiority of the GARCH approach to that of the RiskMetrics model in-sample and within volatility forecasting exercises (see previous chapters), this chapter seeks to answer the same question of superiority within a risk management Value-at-Risk (VaR) scenario. To answer this question daily stock market data from 31 international markets are used in a VaR exceptions forecasting competition. The results show that when forecasting at a 1% VaR the RiskMetrics model does a poor job compared to the GARCH type models, however, at the 5% VaR RiskMetrics does provide an adequate performance.

5.1 Introduction

Despite the empirical success of the GARCH type models within the academic literature the main focus was on finding the best fitting model for stock returns

[36] This chapter is part of a published academic article. Reference: McMillan, D. G., and Kambouroudis, D., (2009), "Are RiskMetrics forecasts good enough? Evidence from 31 stock markets", *International Review of Financial Analysis*, Vol. 18, pp. 117-124.

volatility, its use within the investment community is limited. This, it is suggested, is a result of the complicated nature of the models and that a variety of parameterisations exist, in terms of models that capture asymmetry or long memory or some other nuance of the data. Instead the investment community prefer to adopt a model such as the RiskMetrics exponential smoothing technique, which has the advantage of simplicity over the GARCH models, both in terms of the number of parameters and amount of data required in estimation. The natural question that arises is whether the RiskMetrics approach is in some sense good enough in terms of volatility forecasting and on applications of volatility forecasting. This chapter seeks to examine the volatility forecasts of the RiskMetrics model and a variety of GARCH models for a large selection of international equity markets within a Value-at-Risk (VaR) framework.

Early empirical studies examining the forecasting performance of GARCH models against so-called simple models (including the RiskMetrics approach) often found favour for this latter modelling set, although the whole collection of evidence was mixed. In particular, Figlewski (1997) argued that volatility forecasting models based upon moving averages of historical volatility often provided the best forecasts. Further evidence in support of forecasting models employing simple historical data and against the more involved GARCH modelling approach was provided by Cumby et al. (1993) and Jorion (1995, 1996). However, as noted, evidence was mixed with several papers providing support for the GARCH approach, for example, Brailsford and Faff, 1996; Akigiray, 1989; and McMillan et al. (2000). Furthermore, evidence in favour of the GARCH modelling approach has been strengthened by arguments that

suggest using ex post squared returns as the proxy for true volatility[37] upon which to base comparisons of forecasted volatility is flawed. In particular, Andersen and Bollerslev (1998) and Andersen et al. (1999) have shown that such a measure includes a large noisy component so that volatility forecasts appear artificially poor. Using the realised variance approach further evidence in support of the GARCH approach has been reported by, amongst others, McMillan and Speight (2004), for more on this refer to literature review chapter.

The rest of this chapter is structured as follows; section 2 gives a brief history of Risk Management in Finance and introduces the VaR methodology. Section 3 is the data and methodology section, in section 4 the results with some analysis are presented, finally some concluding remarks are made in section 5.

5.2 Risk Management in finance and VaR[38]

"Optimal risk behaviour takes risks that are worthwhile. This is the central paradigm of finance; we must take risks that are equally rewarded. Both the risks and the rewards are in the future, so it is the expectation of loss that is balanced against the expectation of reward. Thus optimise our behaviour, and in particular our portfolio, to maximise rewards and minimise risks" (p 326, Engle 2003). Risk management is the cornerstone of finance theory. From the seminal work of Markowitz (1952) and Tobin (1958) risk is associated with the variance of the value of a portfolio, minimisation of risk led to portfolio optimisation (and banking behaviour), followed

[37] As a measure of true volatility daily squared returns are used, based on the generally accepted benchmark established by Pagan and Schwert (1990).
[38] The main source for this section is Holton (2002)

by the work of Sharpe (1964) who modelled the relationship between variance (risk) and expected returns with the introduction of the CAPM and in the 1970's Black and Scholes (1972) and Merton (1973) developed the option pricing model which is consistent with the CAPM, Engle (2003).

VaR in its most simple and general form answers the question: 'how much can I afford to lose on this investment'? A question asked by investors and banks. The appealing factor of the VaR technique is that it is a single numerical value of risk measurement for the setting of capital adequacy limits for banks and other institutions. According to city analysts VaR is slowly replacing standard deviation or volatility as the most widely used measure of risk.[39]

The origins of VaR can be traced back as far as 1922 to capital requirements the New York Stock Exchange imposed on member firms and in portfolio theory according to Holton (2002). Holton in his work History of Value-at-Risk 1922-1998 gives a good review and in depth analysis of the VaR technique identifying the basic theories and fundamentals for the development of the technique. Leavens (1945) was the first to ever present a VaR example in his study and although he did not specifically illustrate the VaR metric he mentions: "the spread between probable losses and gains". Both Markowitz (1952) and Roy (1952) independently published a measure of VaR within the portfolio optimization framework. Later Markowitz (1959) in his book 'Portfolio Selection: Efficient Diversification of Investments' proposed alternative calculations for the VaR measure. The VaR measure can then be found in William Sharpe's PhD

[39] Peter Urbani, City Analyst (2002).

thesis and in his 1963 paper 'A simplified model for portfolio analysis' which will prepare the ground for his 1964 Capital Asset Pricing Model (CAPM).

Due to the limited power of processing ability before the 1970's mainly theoretical VaR models were published within the context of portfolio theory, for example Tobin (1958), Treynor (1961), Sharpe (1964), Linter (1965) and Mossin (1966) where VaR measures were best employed for equity portfolios. Later in the 1970's and 1980's changes in the markets such as the collapse of the Bretton Woods agreement in 1971 and the OPEC crises -for more information refer to the work of Holton (2002), innovations in technology and the growth of the financial data industry, the use and implementation of the VaR measure was still a theoretical tool, however markets were becoming more volatile and the need for a measure of financial risk was becoming a necessity.

The 1990's were characterised by large losses and corporation mishaps and the need for a new regulatory framework for acceptable risk management practices. The earlier development by JP Morgan of the RiskMerics helped publicise the VaR to a wide audience. The Basle Committee (1996) updated the previous accord which went into effect in 1998 to include bank capital requirements for market risk. This way financial institutions in addition to the conventional capital requirements for credit risk should also maintain capital against their market VaR.

According to Duffie and Pan (1997), VaR is a standard benchmark for the disclosure of financial risk. As already mentioned, for a given time horizon t and confidence level p, the VaR is the loss in market value over the time horizon t that is expected

with probability $1-p$. Some examples from Duffie and Pan (1997) include the Derivatives Policy Group who has proposed a standard for over-the-counter derivatives broker-dealer reports to the Securities and Exchange Commission that would set a time horizon of two weeks and a confidence level of p 99%. Statistically this VaR measure is the "0.01 critical value" of the probability distribution of changes in market value. Some firms use an overnight VaR measure for internal purposes, as opposed to the two-week standard that is commonly requested for disclosure to regulators, and the 99% confidence level is far from uniformly adopted. For example, J.P. Morgan discloses its daily VaR at the 95% level and Bankers Trust discloses its daily VaR at the 99% level. More specifically, Bankers Trust revealed in its 1994 annual report that its daily VaR was an average of $35 million at the 99% confidence level over one day; this number can be compared to with its annual profit of $615 million or total equity of $4.7 billion. Hence and on the basis of such data different stakeholder groups can decide whether they feel comfortable with the set level of risk (Jorion, 1996).

VaR is a benchmark for relative judgements in sustaining a firm's risk. The risk of one portfolio relative to another, the relative impact on risk of a trade, the modelled risk relative to the historical experience of marks to market, the risk of one volatility environment relative to another (this is what this chapter attempts to do), etc. Even if accurate such comparisons are specific to the time horizon and the confidence level associated with the VaR standard chosen.

Measuring VaR[40]

Once the two quantitative factors of the length of the holding horizon and the confidence level are defined[41], a portfolio's VaR can also be defined. Using the example presented by Jorion (1996), the Basle Committee defined a VaR measure using a 99% confidence interval over 10 trading days. To compute the VaR of a portfolio, define W_0 as the initial investment and R as its rate of return. The portfolio value at the end of the target horizon is $W = W_0(1+R)$. Define μ and σ as the annual mean and standard deviation of R, respectively, and Δt as the time interval considered. If successive returns are uncorrelated, the expected return and risk are then $\mu\Delta t$ and $\alpha\sqrt{\Delta t}$ over the holding horizon. The VaR is defined as the dollar loss relative to what was expected, that is:

$$VaR = E(W) - W^* = W_0(\mu - R^*) \qquad (5.1)$$

Where W^* is the lowest portfolio value at a given confidence level c. Finding VaR is equivalent to identifying the minimum value, W^*, or the cutoff return, R^*. The above definitions are then modified in section 5.3.3, later in the chapter, for the purposes to the volatility forecasting exercise.

[40] Main source for this section is Jorion, (1996).
[41] Both factors (horizon and confidence level) are arbitrary.

5.3 Data and Methodology

5.3.1 Data

In an attempt to build on and progress from the previous two empirical chapters the data for chapter 5 is updated to include more recent observations and more sample countries. The sample consists of 31 international stock market indices. More specifically the G7 consisting of in alphabetical order Canada, France, Germany, Italy, Japan, UK and the USA. Thirteen European markets excluding those already mentioned in the G7 group, namely in alphabetical order Austria, Belgium, Denmark, Finland, Greece, Iceland, Ireland, Netherlands, Portugal, Spain, Sweden, Switzerland, and Turkey. Finally, eleven markets from Asia are included in the sample, Australia, China, Hong Kong, India, Indonesia, Korea, Malaysia, Philippines, Singapore, Taiwan and Thailand. For all the countries daily closing stock prices from the start of 1990 to the end of September 2007 are obtained from Datasream market information service. Price indices are converted to returns by the standard method of calculating the log-differences. As a proxy of true volatility[42] the procedure in Pagan and Schwert (1990) is followed where the squared error from the conditional mean model for returns is used. In the table below the summary statistics of the volatility of the series are presented. The descriptive statistics consist of the arithmetic mean, standard deviation, skewness, kurtosis, minimum and maximum values. The AR column contains the first order autocorrelation coefficient. The asterisk denotes the statistical significance at the 1% level.

[42] As a measure of true volatility daily squared returns are used, based on the generally accepted benchmark established by Pagan and Schwert (1990).

Table 5.1 Descriptive statistics of volatility – G7							
	Mean	SD	Skew	Kurt	Min	Max	AR
G7							
Canada	0.72	2.17	14.36	346.61	<0.01	71.66	0.11*
France	1.66	3.70	6.50	65.76	<0.01	58.95	0.17*
Germany	0.80	2.60	21.93	836.42	<0.01	114.60	0.19*
Italy	1.50	3.34	6.62	72.03	<0.01	59.41	0.20*
Japan	2.02	4.74	11.06	262.80	<0.01	154.51	0.10*
UK	1.00	2.31	6.36	61.17	<0.01	34.84	0.23*
USA	0.95	2.35	8.55	124.83	<0.01	50.59	0.20*

Table 5.2 Descriptive statistics of volatility – Europe							
	Mean	SD	Skew	Kurt	Min	Max	AR
Europe							
Austria	0.60	1.58	16.48	445.31	<0.01	55.48	0.29*
Belgium	1.02	2.97	11.10	215.15	<0.01	87.12	0.32*
Denmark	1.05	2.34	6.17	62.53	<0.01	39.17	0.23*
Finland	3.05	9.98	15.59	390.45	<0.01	302.89	0.14*
Greece	2.65	7.93	11.17	225.79	<0.01	234.43	0.21*
Iceland	0.53	1.56	8.62	113.53	<0.01	31.58	0.16*
Ireland	0.94	2.49	8.52	117.75	<0.01	57.29	0.19*
Netherlands	8.96	50.29	20.65	625.15	<0.01	1841.68	0.06*
Portugal	0.94	3.00	13.44	295.49	<0.01	91.96	0.23*
Spain	1.57	3.76	7.48	91.45	<0.01	78.77	0.20*
Sweden	2.02	4.91	8.73	135.53	<0.01	121.50	0.20*
Switzerland	1.24	3.39	8.63	111.01	<0.01	68.87	0.23*
Turkey	8.37	20.09	7.27	87.15	<0.01	399.14	0.25*

Table 5.3 Descriptive statistics of volatility – Asia							
	Mean	SD	Skew	Kurt	Min	Max	AR
Asia							
Australia	1.32	3.81	9.06	117.67	<0.01	75.68	0.24*
China	6.54	82.52	56.78	3521.05	<0.01	5171.79	0.02
HK	2.33	8.30	18.28	511.61	<0.01	297.46	0.35*
India	2.65	8.12	13.91	347.61	<0.01	276.92	0.23*
Indonesia	2.18	7.73	10.91	167.96	<0.01	172.34	0.18*
Korea	3.38	8.44	7.29	89.33	<0.01	163.96	0.18*
Malaysia	2.14	14.46	26.99	900.09	<0.01	583.39	0.49*
Philippines	2.34	7.71	17.18	476.19	<0.01	261.71	0.11*
Singapore	1.56	5.60	18.57	574.70	<0.01	221.07	0.21*
Taiwan	3.35	8.16	6.25	68.51	<0.01	164.76	0.22*
Thailand	2.92	8.62	10.97	215.86	<0.01	258.03	0.30*

The summary statistics are broadly consistent across all markets and regions, with very few systematic differences between them. In general the mean level of volatility is similar and in the range of 1 to 3, with a few exceptions for example China and the

Netherlands. It appears that the level of volatility is slightly higher in the Asian markets compared to the markets of the G7 and Europe. Likewise, the standard deviations are in a broadly comparable range, with few exceptions noted, and it can again be argued that they are higher in the Asian markets. All the volatility series are characterised by substantial positive skewness and excess kurtosis, reflecting the numerous shocks that affect international equity markets. The maximum values equally reflect the existence of large news events that severely impact on equity markets. The statistical significance of the first order Q-statistic reflects the volatility clustering or serially correlated nature of volatility.[43] Only in the case of China the statistic is not significant however, higher orders lags are significant. Overall, the similarity of these statistics reflects the integrated nature of global equity markets, although the observable higher volatility in Asian markets also reflects their more recent development.

5.3.2 Methodology

The RiskMetrics approach to volatility forecasting uses the simple exponential smoothing model whereby today's volatility forecasts is a weighted average of yesterdays volatility forecast and yesterdays actual volatility:

$$\sigma^2{}_t = \lambda\sigma_{t-1}^2 + (1-\lambda)\, r_{t-1}^2 = (1-\lambda)\sum_{\tau=1}^{\infty} r_{t-\tau}^2 \qquad (5.2)$$

where σ_t^2 is the forecast of volatility and r_t^2 is the squared return, which acts as the proxy for true volatility. Note that through backward substitution of the RiskMetrics

[43] The test performed is the LJung-Box test.

model we arrive at the second expression in equation (5.2) whereby the prediction of volatility is an exponentially weighted moving average of past squared returns. Although in principle the smoothing parameter, λ, can be estimated, the RiskMetrics approach is to fix this value at 0.94 for daily forecasts.

The RiskMetrics approach has some clear advantages. First, it broadly tracks day-to-day volatility changes, whereby recent returns matter more for tomorrow's volatility than distant returns, as λ is less than one. Second, relatively little data needs to be stored. Once a starting value for volatility is found, the only variables needed to calculate tomorrow's volatility is today's volatility and today's squared return, both of which are known at the end of trading today. Third, the model only contains one unknown parameter, which, as noted, is typically set to $\lambda = 0.94$, hence, no estimation is necessary, which is a huge advantage in large portfolios. However, the RiskMetrics approach is also subject to two main shortcomings, first it is not able to capture the asymmetry effects often noted in equity data, that is, the negative correlation between returns and volatility. Second, that the model is not able to provide long-horizon forecasts. Most time-series models, such as GARCH type models, will have forecasts that tend towards the unconditional variance of the series as the prediction horizon increases. This is a good property for a volatility forecasting models to have, since it is well known that volatility series are 'mean reverting'. This feature is accounted for in GARCH volatility forecasting models but not by EWMA, Brooks (2008).

The GARCH approach to volatility forecasting is more involved than the RiskMetrics model and extracts the conditional variance from the returns series:

$$r_t = \varepsilon_t = \sigma_t z_t \tag{5.3}$$

with z_t ~ iid N(0, 1)

And the GARCH(1,1) model then written as:

$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{5.4}$$

The GARCH model thus contains three components, the constant that represents mean volatility, news about volatility from the previous period measured by the lagged squared return and last period's forecast variance.

The RiskMetrics model can now be viewed as a special case of the GARCH model where $\alpha_1 = 1 - \lambda$, $\beta = \lambda$ (hence note that $\alpha + \beta = 1$) and $\alpha_0 = 0$. This is a key difference, for a finite unconditional to exist $\alpha + \beta < 1$, in which case $\sigma^2 = \alpha_0/(1-\alpha_1-\beta)$. It is now clear that in the RiskMetrics model the long-run variance is infinite or is not well-defined. Thus, an important quirk of the RiskMetrics model is that it ignores the fact that the long-run average variance tends to be relatively stable over time.

The GARCH model, in turn, implicitly relies on σ^2. This can be seen by solving for α_0 in the long-run variance equation and substituting it into the dynamic variance equation as such:

$$\sigma_t^2 = \sigma^2 + \alpha_1 (r_{t-1}^2 - \sigma^2) + \beta(\sigma_{t-1}^2 - \sigma^2) \tag{5.5}$$

Thus, current variance is the long-run average variance with something added (subtracted) if yesterday's squared news is above (below) its long-run average, and something added (subtracted) if yesterday's variance is above (below) its long-run average. To further highlight this property we can again obtain the *k*-step ahead forecast:

$$\sigma^2_{t+k|t} = \sigma^2 + (\alpha_1 + \beta)^{k-1}(\sigma^2_{t+1} - \sigma^2) \tag{5.6}$$

While the *k*-step ahead cumulative forecasts comparable to that in equation (5.2) for the RiskMetrics model is given by:

$$\sigma^2_{t\to t+k|t} = k\sigma^2 + (\sigma^2_{t+1} - \sigma^2)(1-(\alpha_1+\beta)^k)(1-\alpha_1-\beta)^{-1} \tag{5.7}$$

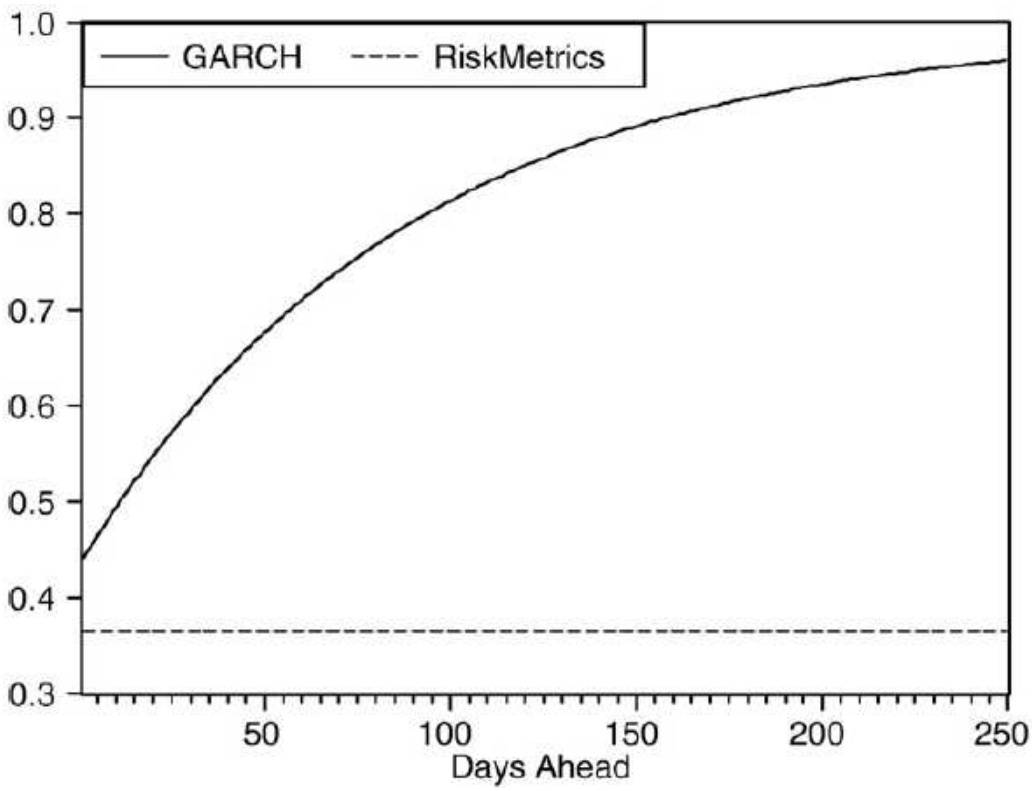


Figure 5.1: GARCH and RiskMetrics multi-step forecasts

Graphically, Figure 5.1 illustrates the difference from a low starting point for volatility, we can see the GARCH forecasts increasing as we forecast further into the future (and ultimately converging with the unconditional variance) whereas the RiskMetrics forecasts remain constant.

The other advantage of the GARCH approach to volatility forecasting over the RiskMetrics model is the flexibility to capture different nuances of the data, including asymmetries between negative and positive shocks and long memory effects. To this end in addition to the GARCH(1,1) model described above we also consider a further selection of models from the GARCH genre as noted below. The discussion of these models is kept brief, but extensive reviews have been provided by, for example, Bollerslev et al. (1992) and Bera and Higgins (1993).

In this section a small reminder of the models to be used in this chapter are presented. These models have extensively been looked into in literature review section 2.6 (page 42) and in Chapter 3 section 3.3.2 (page 69).

The GARCH(1,1) model of Engle (1982) and Bolerslev (1986) using the standard notation is given by:

$$h_{t+1}^2 = \omega + \alpha\varepsilon_t^2 + \beta h_t^2 \qquad (5.8)$$

The EGARCH model of Nelson (1991) is given by:

$$\log(h_{t+1}^2) = \omega + \alpha\left|\frac{\varepsilon_t}{h_t}\right| + \gamma\frac{\varepsilon_t}{h_t} + \beta\log(h_t^2) \tag{5.9}$$

The asymmetric power-ARCH (APARCH) by Ding et al. (1993) is specified as:

$$h_t^\delta = \omega + \alpha_1(|\varepsilon_{t-1}| - \gamma\varepsilon_{t-1})^\delta + \beta_1 h_{t-1}^\delta \tag{5.10}$$

The HYGARCH model by Davidson (2004) is given by:

$$h_t^2 = \omega + [1-\beta(L)]^{-1} + \{1-[1-\beta(L)]^{-1}\phi(L)[1+\alpha((1-L)^d - 1)]\}\varepsilon_t \tag{5.11}$$

The FIGARCH model by Baillie et al. (1996) is given by:

$$h_t^2 = \omega + [1-\beta(L)]^{-1} + \{1-[1-\beta(L)]^{-1}\phi(L)(1-L)^d\}\varepsilon_t \tag{5.12}$$

The IGARCH is specified as an extension to GARCH model for which the following condition must hold, $\alpha + \beta = 1$ for the conditional variance to be clearly non-stationary.

5.3.3 Methodology for comparisons of forecast performance

The VaR of a portfolio is calculated as:

$$VaR = \alpha(N)\sigma_{t+1}V \tag{5.13}$$

where $\alpha(N)$ defines the appropriate left-hand cut-off of the normal distribution, σ_{t+1} is the one-step ahead volatility forecast and V refers to the value of the portfolio, converting VaR estimates into currency value.

For the analysis of the performance of the volatility forecasting models for producing 'good' VaR estimates is achieved by reporting how many times the actual loss exceeds the VaR estimation. In this chapter the Kupiec (1995) and the Dynamic Quantile (DQ) by Engle and Manganelli are testing procedures are used.

One popular test of VaR accuracy was proposed by Kupiec (1995). This test belongs in the category of tests known as VaR backtests and is concerned with whether or not the reported VaR is violated more (or less) than α * 100% of the time. The proportion of failures examines how many times the VaR is violated over a given time span and if the number of violations differs considerably from α * 100% of the sample, then the accuracy of the VaR model is questioned, Campbell (2005).

The Kupiec (1995) test is computed for testing the equality of the frequency of exceptions and the chosen left-hand tail cut-off. The test defines a likelihood ratio (LR) test statistic as:

$$LR = -2\log\left(\frac{\alpha^N(1-\alpha)^{T-N}}{f^N(1-f)^{T-N}}\right) \tag{5.14}$$

where N is the number of VaR violations, T is the total number of observations and α is the theoretical failure rate. Under the null hypothesis that *f* is the failure rate, the LR test statistic is asymptotically distributed as a chi square distribution $X^2(1)$.[44]

Argued by Engle and Manganelli the Kupiec test is an unconditional test of VaR accuracy and besides failure rate a relevant VaR model should feature a sequence of indicator functions that are not serially correlated. Given the substantial time-variation within volatility, conditional accuracy of the VaR estimates are also important, to this Engle and Manganelli (2004) proposed the test of Dynamic Quantile. For a correctly specified VaR model, not only should the exceptions occur at the specific rate (1% or 5%) but also should be independent and identically distributed.

Engle and Magnanelli (2004/1999) Dynamic Quantile test is defined by the sequence:

$$Hit_k = I(r_k < -VaR_k) - \alpha \qquad (5.15)$$

The sequence assumes the value $(1-\alpha)$ when the returns r_t are less than the VaR quantile and the value of $-\alpha$ otherwise, with the expected value of Hit_k equal to zero. This sequence should then be uncorrelated with past information and a mean value of zero where there will be no autocorrelation in the hits and there will be the correct fraction of exceptions –this property is tested by the Kupiec test. To test for autocorrelation in the hit sequence Hit_k is regressed on five lags (five days) and the current value of VaR. The DQ test statistic is computed as:

[44] G@rch 6 Help http://www.core.ucl.ac.be/~laurent/G@RCH/site/default.htm

$$DQ = \frac{\hat{\beta}'x'x\hat{\beta}}{\alpha(1-\alpha)} \tag{5.16}$$

Where χ is the vector of explanatory variables and $\hat{\beta}$ the OLS estimates. The DQ test distributed as a X^2 distribution with degrees of freedom equal to the number of parameters.

5.4 Results and analysis

The purpose of this exercise is to determine which forecast model provides the best VaR estimates and this is done by taking into account practices used within the regulatory environment. For this reason daily VaR measures are produced while the in-sample and out-of-sample dates are recursively updatde every sixty days (the 60 day window is approximately 3 months of trading often regarded as the length of time over which practitioners evaluate their models) in addition both 1% and 5% VaR's are calculated for each market.

In the tables 5.4 - 5.6.b the results of the VaR exercise for the 1% failure rate are shown. For each model and for each sample country three statistics are presented; the Failure Rate for which the highest value indicates the least accurate model, the Kupiec Test whose values are compared to the critical value of $X^2(1)$ determining the significance of the model and in the same way the Dynamic Quantile Testis also performed. These results are then followed by summary tables providing a more comprehensive overview of the results.

Table 5.4 Results for 1% VaR Failure Rate - G7

	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test
Model	Canada			France			Germany			Italy		
RM	0.019	8.23 (0.00)	28.16 (0.00)	0.020	9.58 (0.00)	35.42 (0.00)	0.021	12.52 (0.00)	36.58 (0.00)	0.027	25.07 (0.00)	66.24 (0.00)
G	0.017	4.71 (0.03)	24.25 (0.00)	0.011	0.15 (0.70)	6.49 (0.48)	0.018	6.97 (0.01)	23.97 (0.00)	0.016	3.73 (0.05)	25.43 (0.00)
E	0.017	4.71 (0.03)	25.51 (0.00)	0.010	0.01 (0.91)	13.02 (0.07)	0.020	9.58 (0.00)	43.45 (0.00)	0.017	4.71 (0.03)	27.62 (0.00)
AP	0.018	6.97 (0.01)	27.61 (0.00)	0.009	0.21 (0.64)	8.04 (0.33)	0.019	8.23 (0.00)	26.32 (0.00)	0.016	3.73 (0.05)	23.98 (0.00)
IG	0.016	3.73 (0.05)	25.96 (0.00)	0.013	1.40 (0.24)	11.72 (0.11)	0.016	3.73 (0.05)	21.36 (0.00)	0.020	9.58 (0.00)	37.58 (0.00)
FI	0.016	3.73 (0.05)	13.04 (0.07)	0.011	0.15 (0.70)	13.40 (0.06)	0.018	6.97 (0.01)	25.19 (0.00)	0.022	14.11 (0.00)	32.51 (0.00)
HY	0.015	2.84 (0.09)	9.42 (0.23)	0.011	0.15 (0.70)	13.52 (0.06)	0.019	8.23 (0.00)	34.67 (0.00)	0.022	14.11 (0.00)	34.07 (0.00)
	Japan			UK			USA					
RM	0.022	14.11 (0.00)	64.66 (0.00)	0.021	12.5 (0.00)	37.16 (0.00)	0.021	12.52 (0.00)	45.38 (0.00)			
G	0.013	1.40 (0.24)	10.38 (0.17)	0.015	3.73 (0.05)	13.30 (0.07)	0.019	8.23 (0.00)	41.79 (0.00)			
E	0.014	2.06 (0.15)	10.19 (0.18)	0.016	3.73 (0.05)	13.83 (0.05)	0.013	1.40 (0.24)	24.95 (0.00)			
AP	0.013	1.40 (0.24)	7.30 (0.40)	0.015	2.84 (0.09)	11.43 (0.12)	0.013	1.40 (0.24)	10.05 (0.19)			
IG	0.013	0.85 (0.36)	11.33 (0.12)	0.018	6.97 (0.01)	24.74 (0.00)	0.019	8.23 (0.00)	42.97 (0.00)			
FI	0.014	2.06 (0.15)	14.49 (0.04)	0.013	1.40 (0.24)	13.17 (0.07)	0.016	3.73 (0.05)	10.70 (0.15)			
HY	0.014	2.06 (0.15)	14.48 (0.04)	0.013	1.40 (0.24)	13.18 (0.07)	0.014	2.06 (0.15)	12.21 (0.09)			

Table 5.5.a Results for 1% VaR Failure Rate – Europe

	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test
Model	Austria			Belgium			Denmark			Finland		
RM	0.026	23.08 (0.00)	59.12 (0.00)	0.021	11.01 (0.00)	76.62 (0.00)	0.020	9.58 (0.00)	39.01 (0.00)	0.018	6.97 (0.01)	43.48 (0.00)
G	0.015	2.84 (0.09)	49.74 (0.00)	0.017	5.79 (0.02)	5.04 (0.00)	0.018	6.97 (0.01)	25.04 (0.00)	0.013	1.40 (0.24)	24.57 (0.00)
E	0.015	2.84 (0.09)	49.41 (0.00)	0.016	3.73 (0.05)	39.83 (0.00)	0.018	6.97 (0.01)	33.29 (0.00)	0.017	4.71 (0.03)	25.26 (0.00)
AP	0.013	1.40 (0.24)	26.86 (0.00)	0.014	2.06 (0.15)	26.68 (0.00)	0.017	5.79 (0.02)	25.41 (0.00)	0.014	2.06 (0.15)	22.93 (0.00)
IG	0.013	1.40 (0.24)	26.06 (0.00)	0.014	2.06 (0.15)	69.69 (0.00)	0.017	5.79 (0.02)	25.70 (0.00)	0.013	0.85 (0.36)	25.35 (0.00)
FI	0.021	11.01 (0.00)	79.11 (0.00)	0.013	0.85 (0.36)	22.94 (0.00)	0.018	6.97 (0.01)	25.04 (0.00)	0.013	1.40 (0.24)	24.04 (0.00)
HY	0.024	17.49 (0.00)	75.82 (0.00)	0.016	3.73 (0.05)	40.09 (0.00)	0.018	6.97 (0.01)	25.04 (0.00)	0.013	1.40 (0.24)	24.13 (0.00)
	Greece			Iceland			Ireland			Netherlands		
RM	0.021	11.01 (0.00)	80.45 (0.00)	0.017	5.79 (0.02)	26.04 (0.00)	0.021	11.01 (0.00)	44.79 (0.00)	0.023	15.77 (0.00)	82.79 (0.00)
G	0.011	0.15 (0.70)	26.49 (0.00)	0.017	4.71 (0.03)	27.69 (0.00)	0.017	4.71 (0.03)	10.10 (0.18)	0.021	12.52 (0.00)	82.41 (0.00)
E	0.013	0.85 (0.36)	54.73 (0.00)	0.018	6.97 (0.01)	68.49 (0.00)	0.015	2.84 (0.09)	5.69 (0.58)	0.021	15.52 (0.00)	81.74 (0.00)
AP	0.010	0.01 (0.91)	28.87 (0.00)	0.018	6.97 (0.01)	48.92 (0.00)	0.016	3.73 (0.05)	7.25 (0.40)	0.011	0.15 (0.70)	54.42 (0.00)
IG	0.011	0.15 (0.70)	26.50 (0.00)	0.014	2.06 (0.15)	25.67 (0.00)	0.015	2.84 (0.09)	5.72 (0.57)	0.047	94.30 (0.00)	87.54 (0.00)
FI	0.014	2.06 (0.15)	25.54 (0.00)	0.014	2.06 (0.15)	42.24 (0.00)	0.015	2.84 (0.09)	10.44 (0.16)	0.022	17.55 (0.00)	80.56 (0.00)
HY	0.014	2.06 (0.15)	25.63 (0.00)	0.018	6.97 (0.01)	45.57 (0.00)	0.015	2.84 (0.09)	10.42 (0.17)	0.020	9.58 (0.00)	56.29 (0.00)

Table 5.5.b Results for 1% VaR Failure Rate – Europe

	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test
Model	Portugal			Spain			Sweden			Switzerland		
RM	0.018	6.97 (0.01)	25.10 (0.00)	0.24	17.49 (0.00)	41.62 (0.00)	0.025	19.29 (0.00)	48.40 (0.00)	0.021	12.52 (0.00)	32.73 (0.00)
G	0.013	1.40 (0.24)	24.64 (0.00)	0.014	2.06 (0.15)	25.65 (0.00)	0.017	4.71 (0.03)	25.45 (0.00)	0.014	2.06 (0.15)	7.10 (0.42)
E	0.013	0.85 (0.36)	25.56 (0.00)	0.015	2.84 (0.09)	24.50 (0.00)	0.013	0.85 (0.36)	28.61 (0.00)	0.011	0.15 (0.70)	6.62 (0.47)
AP	0.013	0.85 (0.36)	10.01 (0.19)	0.015	2.84 (0.09)	24.52 (0.00)	0.013	1.40 (0.24)	27.11 (0.00)	0.013	0.85 (0.36)	6.33 (0.50)
IG	0.013	1.40 (0.24)	24.74 (0.00)	0.015	2.84 (0.09)	25.11 (0.00)	0.016	3.73 (0.05)	26.76 (0.00)	0.011	0.15 (0.70)	7.46 (0.38)
FI	0.013	1.40 (0.24)	25.00 (0.00)	0.017	4.71 (0.03)	26.59 (0.00)	0.015	2.84 (0.09)	25.45 (0.00)	0.013	1.40 (0.24)	6.28 (0.51)
HY	0.013	1.40 (0.24)	24.10 (0.00)	0.016	3.73 (0.05)	26.95 (0.00)	0.015	2.84 (0.09)	25.40 (0.00)	0.014	2.06 (0.15)	7.10 (0.42)
	Turkey											
RM	0.017	4.71 (0.03)	11.92 (0.10)									
G	0.009	0.21 (0.64)	0.83 (0.99)									
E	0.017	4.71 (0.03)	13.75 (0.06)									
AP	0.008	0.58 (0.45)	0.97 (0.99)									
IG	0.010	0.01 (0.91)	0.96 (0.99)									
FI	0.013	0.85 (0.36)	2.38 (0.94)									
HY	0.013	1.40 (0.24)	6.70 (0.46)									

Table 5.6.a Results for 1% VaR Failure Rates – Asia

	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test
Model	Australia			China			Hong Kong		
RM	0.018	6.97 (0.01)	51.44 (0.00)	0.015	2.84 (0.09)	10.62 (0.16)	0.022	14.11 (0.00)	40.76 (0.00)
G	0.013	0.85 (0.36)	26.78 (0.00)	0.011	0.15 (0.70)	7.71 (0.36)	0.010	0.01 (0.91)	13.61 (0.06)
E	0.017	5.79 (0.02)	58.98 (0.00)	0.005	4.33 (0.04)	19.76 (0.01)	0.010	0.03 (0.86)	7.90 (0.34)
AP	0.013	0.85 (0.36)	26.76 (0.00)	0.014	2.06 (0.15)	10.90 (0.14)	0.010	0.03 (0.86)	7.80 (0.35)
IG	0.012	0.44 (0.51)	28.76 (0.00)	0.013	0.85 (0.36)	7.66 (0.36)	0.011	0.15 (0.70)	13.06 (0.07)
FI	0.017	4.71 (0.03)	80.46 (0.00)	0.019	8.23 (0.00)	16.77 (0.02)	0.014	2.06 (0.15)	31.79 (0.00)
HY	0.017	4.71 (0.03)	80.49 (0.00)	0.020	9.58 (0.00)	21.81 (0.00)	0.014	2.06 (0.15)	31.97 (0.00)
	India			Indonesia			Korea		
RM	0.025	21.15 (0.00)	101.57 (0.00)	0.020	9.58 (0.00)	106.59 (0.00)	0.025	19.29 (0.00)	56.94 (0.00)
G	0.017	4.71 (0.03)	45.65 (0.00)	0.016	3.73 (0.05)	35.63 (0.00)	0.017	4.71 (0.03)	25.90 (0.00)
E	0.017	4.71 (0.03)	46.09 (0.00)	0.017	4.71 (0.03)	54.38 (0.00)	0.019	8.23 (0.00)	29.72 (0.00)
AP	0.017	4.71 (0.03)	45.55 (0.00)	0.016	3.73 (0.05)	35.68 (0.00)	0.017	5.79 (0.02)	25.83 (0.00)
IG	0.017	4.71 (0.03)	45.69 (0.00)	0.016	3.73 (0.05)	35.68 (0.00)	0.017	4.71 (0.03)	27.10 (0.00)
FI	0.018	6.97 (0.01)	42.96 (0.00)	0.020	9.58 (0.00)	51.19 (0.00)	0.017	4.71 (0.03)	25.79 (0.00)
HY	0.017	5.79 (0.02)	44.23 (0.00)	0.018	6.97 (0.01)	33.95 (0.00)	0.017	4.71 (0.03)	27.50 (0.00)

Table 5.6.b Results Table 1% VaR Failure Rates – Asia

	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test
Model	Malaysia			Philippines			Singapore		
RM	0.021	11.01 (0.00)	42.14 (0.00)	0.017	5.79 (0.02)	12.94 (0.07)	0.025	21.15 (0.00)	55.05 (0.00)
G	0.012	0.44 (0.51)	64.61 (0.00)	0.009	0.21 (0.64)	8.51 (0.29)	0.015	2.84 (0.09)	10.63 (0.16)
E	0.014	2.06 (0.15)	52.49 (0.00)	0.010	0.03 (0.86)	8.94 (0.26)	0.014	2.06 (0.15)	10.73 (0.15)
AP	0.013	0.85 (0.36)	59.79 (0.00)	0.010	0.03 (0.86)	1.74 (0.97)	0.013	1.40 (0.24)	10.36 (0.17)
IG	0.013	0.85 (0.36)	60.09 (0.00)	0.009	0.21 (0.64)	1.00 (0.99)	0.014	2.06 (0.15)	10.57 (0.16)
FI	0.017	5.79 (0.02)	73.15 (0.00)	0.013	0.85 (0.36)	6.88 (0.44)	0.020	9.58 (0.00)	26.74 (0.00)
HY	0.017	4.71 (0.03)	47.34 (0.00)	0.013	0.85 (0.36)	6.80 (0.45)	0.020	9.58 (0.00)	26.71 (0.00)
	Taiwan			Thailand					
RM	0.021	12.52 (0.00)	50.52 (0.00)	0.017	4.71 (0.03)	48.57 (0.00)			
G	0.015	2.84 (0.09)	48.56 (0.00)	0.010	0.03 (0.86)	31.74 (0.00)			
E	0.014	2.06 (0.15)	51.56 (0.00)	0.010	0.03 (0.86)	31.56 (0.00)			
AP	0.013	1.40 (0.24)	55.67 (0.00)	0.009	0.21 (0.64)	9.01 (0.25)			
IG	0.014	2.06 (0.15)	51.48 (0.00)	0.009	0.21 (0.64)	35.36 (0.00)			
FI	0.17	4.71 (0.03)	45.02 (0.00)	0.011	0.15 (0.70)	61.67 (0.00)			
HY	0.017	5.79 (0.02)	44.30 (0.00)	0.011	0.15 (0.70)	61.54 (0.00)			

In the summary table 5.7 below the VaR exercise is carried out for the 1% failure rate. The RiskMetrics model for all the G7 countries has the highest failure rate suggesting that the more advanced GARCH type models do a better overall job. This result can also be confirmed if average failure rates are considered. Specifically for the G7 the RiskMetrics models has an average failure rate of 0.021 whereas the rest of the models are between 0.015 and 0.016. The same conclusion is drawn when the the results of the Kupiec test and the Dynamic Quantile are considered. The Kupiec test statistic is found significant for all the G7 countries as is the Dynamic Quantile test statistic. These results suggests that in both cases the null hypothesis of a well specified RiskMetrics model is rejected for all the sample countries. The same procedure is followed for the rest of the models where only between two or three countries the Kupiec test returns a significant result and for the Dynamic Quantile test between three and five countries give a significant result. In trying to identify a winner for the G7 sample, it is not straight forward but the APARCH and HYGARCH models seem to dominate with the lowest Failure rate and with the least number of significant test statistics, however for the worst performing model the answer is a simple one and this is the RiskMetrics model.

Table 5.7 Summary of 1% VaR failure tests - G7

Model	Average Failure Rate	Significant Kupiec Test	Significant DQ test
RiskMetrics	0.021	All	All
GARCH	0.016	Canada, Germany, USA	Canada, Germany, Italy, USA
EGARCH	0.015	Canada, Germany, Italy	Canada, Germany, Italy, USA
APARCH	0.015	Canada, Germany	Canada, Germany, Italy
IGRACH	0.016	Italy, UK, USA	Canada, Germany, Italy, USA, UK
FIGARCH	0.016	Germany, Italy	Germany, Italy, Japan, USA
HYGARCH	0.015	Germany, Italy	Germany, Italy, Japan

In summary table 5.8 the same as in table 5.7 procedure is followed but for a sample of 13 European markets. The results are similar to those reported for table 5.7. The RiskMetrics model has the highest number of cases with the maximum Failure rate. The average failure rate for the RiskMetrics model is 0.021 and for the rest of the models it is between 0.013 and 0.016. In looking at the results of the Kupiec test and the Dynamic Quantile again for all the European countries the RiskMetrics returns in every case significant test statistics indicating that it is comparatively a weaker model. The number of markets with a significant Kupiec test result for the GARCH type models is between two and six, but for the Dynamic Quantile the range is between nine and ten markets. As before clearly the worst performer is the RiskMetrics model and the best the APARCH and IGARCH models.

Table 5.8 Summary of 1% VaR failure tests – Europe

Model	Average Failure Rate	Significant Kupiec Test	Significant DQ test
RiskMetrics	0.021	All	All
GARCH	0.015	Belgium, Denmark, Iceland, Ireland, Netherlands, Sweden	Austria, Belgium, Denmark, Finland, Greece, Iceland, Netherlands, Portugal, Spain, Sweden
EGARCH	0.016	Denmark, Finland, Iceland, Netherlands, Turkey	Austria, Belgium, Denmark, Finland, Greece, Iceland, Netherlands, Portugal, Spain, Sweden
APARCH	0.013	Denmark, Iceland	Austria, Belgium, Denmark, Finland, Greece, Iceland, Netherlands, Spain, Sweden
IGRACH	0.016	Denmark, Netherlands	Austria, Belgium, Denmark, Finland, Greece, Iceland, Netherlands, Portugal, Spain, Sweden
FIGARCH	0.015	Austria, Denmark, Netherlands, Spain	Austria, Belgium, Denmark, Finland, Greece, Iceland, Netherlands, Portugal, Spain, Sweden
HYGARCH	0.016	Austria, Denmark, Iceland, Netherlands	Austria, Belgium, Denmark, Finland, Greece, Iceland, Netherlands, Portugal, Spain, Sweden

Not a very different picture from before is given in summary table 5.9 where within a 1% VaR setting several models from the AGRCH family are compared to the RiskMetrics model. Looking again at the number of times RiskMetrics gave the maximum Failure rate indicates that also for the 11 Asian markets it is the least accurate model. The average failure rate confirms this as again and average of 0.021 is compared to an average between 0.013 and 0.017. The significance of the Kupiec test and the Dynamic Quantile test although do not give a clear picture with regard the worst performing model with several countries returning significant tests nevertheless the APARCH and IGARCH dominate.

Table 5.9 Summary of 1% VaR failure tests – Asia

Model	Average Failure Rate	Significant Kupiec Test	Significant DQ test
RiskMetrics	0.021	Australia, Hong Kong, India, Indonesia, Korea, Malaysia, Philippines, Singapore, Taiwan, Thailand	Australia, Hong Kong, Indonesia, Korea, Malaysia, Singapore, Taiwan, Thailand
GARCH	0.013	India, Korea	Australia, India, Indonesia, Korea, Malaysia, Taiwan, Thailand
EGARCH	0.013	Australia, China, India, Indonesia, Korea	Australia, China, India, Indonesia, Korea, Malaysia, Taiwan, Thailand
APARCH	0.013	India, Korea	Australia, India, Indonesia, Korea, Malaysia, Taiwan
IGRACH	0.013	India, Korea	Australia, India, Indonesia, Korea, Taiwan, Thailand
FIGARCH	0.017	Australia, China, India, Indonesia, Korea, Malaysia, Singapore, Taiwan	Australia, China, Hong Kong, India, Indonesia, Korea, Malaysia, Singapore, Taiwan, Thailand
HYGARCH	0.016	Australia, China, India, Indonesia, Korea, Malaysia, Singapore, Taiwan	Australia, China, Hong Kong, India, Indonesia, Korea, Malaysia, Singapore, Taiwan, Thailand

Overall we can see that at the 1% VaR, for the G7 several of the GARCH models produce accurate VaR forecasts with insignificant Kupiec and DQ tests. Notably for Canada, France, Japan, the UK and USA where at least one GARCH model performs well, including the APARCH model for four of the series. Examining the European markets, for the majority of the markets at least one GARCH model including the APARCH model (except Iceland) achieves an insignificant Kupiec test (i.e. has the correct frequency of exceptions) if not the DQ test (autocorrelation in the exceptions). Exceptions to this include Denmark for which no model performs well, and Ireland, Switzerland and Turkey for which several models achieve insignificance on both specification tests. The results for the Asian markets are broadly similar in character to those of the European markets. Specifically, for four markets at least one GARCH model, including the APARCH model, achieves insignificance on both tests supporting the adequacy of the model, while for the remaining five series at least one model has an insignificant Kupiec test. Next the 5% VaR results are presented.

Table 5.10 Results for 5% VaR Failure Rate - G7

	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test
Model	Canada			France			Germany			Italy		
RM	0.053	0.26 (0.61)	13.09 (0.07)	0.056	0.79 (0.37)	9.66 (0.21)	0.057	1.30 (0.25)	7.79 (0.35)	0.054	0.41 (0.52)	6.52 (0.48)
G	0.050	0.01 (0.99)	12.18 (0.09)	0.043	1.42 (0.23)	5.47 (0.60)	0.058	1.59 (0.21)	13.62 (0.06)	0.043	1.42 (0.23)	10.38 (0.17)
E	0.049	0.02 (0.90)	7.75 (0.36)	0.044	0.85 (0.36)	8.48 (0.29)	0.062	3.51 (0.06)	14.31 (0.05)	0.043	1.42 (0.23)	8.64 (0.28)
AP	0.051	0.02 (0.90)	10.12 (0.18)	0.044	0.85 (0.36)	6.97 (0.43)	0.059	1.92 (0.17)	3.63 (0.82)	0.044	1.12 (0.29)	8.19 (0.32)
IG	0.047	0.27 (0.60)	15.79 (0.03)	0.044	1.12 (0.29)	4.24 (0.75)	0.045	0.62 (0.43)	3.33 (0.85)	0.045	0.62 (0.43)	8.87 (0.26)
FI	0.050	0.01 (0.99)	11.07 (0.14)	0.45	0.62 (0.43)	13.79 (0.06)	0.051	0.02 (0.90)	10.72 (0.15)	0.049	0.02 (0.90)	8.69 (0.28)
HY	0.050	0.01 (0.99)	11.48 (0.12)	0.043	1.42 (0.23)	8.51 (0.29)	0.056	1.03 (0.31)	5.47 (0.60)	0.05	0.00 (1.00)	8.15 (0.32)
	Japan			UK			USA					
RM	0.056	1.03 (0.31)	15.85 (0.03)	0.060	2.27 (0.13)	7.11 (0.44)	0.60	2.27 (0.13)	16.99 (0.02)			
G	0.044	0.85 (0.36)	6.88 (0.44)	0.051	0.02 (0.90)	4.64 (0.70)	0.048	0.15 (0.70)	8.25 (0.31)			
E	0.043	1.42 (0.23)	8.27 (0.31)	0.051	0.02 (0.90)	1.38 (0.99)	0.048	0.15 (0.70)	8.32 (0.31)			
AP	0.045	0.62 (0.43)	8.61 (0.28)	0.049	0.02 (0.90)	2.58 (0.92)	0.047	0.27 (0.60)	7.20 (0.41)			
IG	0.042	1.76 (0.18)	6.86 (0.44)	0.052	0.07 (0.80)	3.10 (0.88)	0.046	0.43 (0.51)	7.39 (0.39)			
FI	0.044	0.85 (0.36)	8.54 (0.29)	0.044	0.85 (0.36)	3.37 (0.85)	0.044	0.85 (0.35)	7.70 (0.36)			
HY	0.044	0.85 (0.36)	11.62 (0.11)	0.045	0.62 (0.43)	3.28 (0.86)	0.45	0.62 (0.43)	10.61 (0.16)			

Table 5.11.a Results for 5% VaR Failure Rate – Europe

	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test
Model	Austria			Belgium			Denmark			Finland		
RM	0.060	2.27 (0.13)	21.58 (0.00)	0.051	0.02 (0.90)	5.43 (0.61)	0.052	0.15 (0.70)	16.63 (0.02)	0.055	0.58 (0.44)	8.84 (0.26)
G	0.039	3.53 (0.06)	13.79 (0.06)	0.041	2.14 (0.14)	8.18 (0.32)	0.050	0.01 (0.99)	3.90 (0.79)	0.041	2.14 (0.14)	10.86 (0.14)
E	0.044	1.12 (0.29)	9.18 (0.24)	0.040	3.03 (0.08)	8.06 (0.33)	0.046	0.43 (0.51)	2.74 (0.91)	0.040	3.03 (0.08)	11.12 (0.13)
AP	0.044	0.85 (0.36)	8.13 (0.32)	0.044	1.12 (0.29)	5.52 (0.60)	0.050	0.01 (0.99)	4.67 (0.70)	0.41	2.14 (0.14)	9.05 (0.25)
IG	0.043	1.42 (0.23)	8.31 (0.31)	0.035	6.71 (0.01)	13.21 (0.07)	0.042	1.76 (0.18)	3.40 (0.85)	0.040	2.57 (0.11)	11.98 (0.10)
FI	0.048	0.07 (0.79)	6.30 (0.51)	0.037	4.67 (0.03)	9.69 (0.21)	0.048	0.07 (0.79)	3.39 (0.85)	0.037	4.67 (0.03)	8.61 (0.28)
HY	0.051	0.02 (0.90)	12.30 (0.09)	0.040	2.57 (0.11)	8.34 (0.30)	0.048	0.07 (0.79)	3.39 (0.85)	0.037	4.67 (0.03)	8.62 (0.28)
	Greece			Iceland			Ireland			Netherlands		
RM	0.053	0.26 (0.61)	17.40 (0.02)	0.053	0.26 (0.61)	5.68 (0.58)	0.050	0.01 (0.99)	5.53 (0.60)	0.052	0.07 (0.80)	16.02 (0.02)
G	0.035	6.71 (0.01)	19.67 (0.01)	0.048	0.15 (0.70)	14.17 (0.05)	0.052	0.07 (0.80)	7.87 (0.34)	0.041	1.76 (0.18)	16.28 (0.02)
E	0.037	4.67 (0.03)	27.78 (0.00)	0.050	0.01 (0.95)	24.76 (0.01)	0.050	0.01 (0.99)	4.27 (0.75)	0.030	12.80 (0.00)	35.12 (0.00)
AP	0.034	7.49 (0.01)	21.04 (0.00)	0.048	0.15 (0.70)	13.93 (0.05)	0.048	0.07 (0.80)	5.81 (0.56)	0.022	25.60 (0.00)	16.10 (0.02)
IG	0.034	7.49 (0.01)	21.32 (0.00)	0.042	1.76 (0.18)	9.93 (0.19)	0.043	1.42 (0.23)	4.77 (0.69)	0.068	7.97 (0.00)	30.28 (0.00)
FI	0.041	2.14 (0.14)	15.07 (0.04)	0.049	0.02 (0.90)	22.64 (0.00)	0.050	0.01 (0.99)	4.96 (0.66)	0.038	4.66 (0.03)	19.46 (0.00)
HY	0.043	1.42 (0.23)	14.94 (0.04)	0.048	0.07 (0.79)	23.53 (0.00)	0.049	0.02 (0.90)	5.34 (0.62)	0.040	2.57 (0.11)	13.43 (0.06)

Table 5.11.b Results for 5% VaR Failure Rate – Europe

	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test
Model	Portugal			Spain			Sweden			Switzerland		
RM	0.052	0.15 (0.70)	7.75 (0.36)	0.052	0.07 (0.80)	9.48 (0.22)	0.063	3.97 (0.05)	20.20 (0.01)	0.063	3.97 (0.05)	14.64 (0.04)
G	0.044	1.12 (0.29)	3.32 (0.85)	0.034	7.49 (0.01)	9.85 (0.20)	0.048	0.07 (0.80)	0.81 (0.99)	0.049	0.02 (0.90)	3.74 (0.81)
E	0.042	1.76 (0.18)	3.59 (0.83)	0.037	4.67 (0.03)	6.84 (0.45)	0.049	0.02 (0.90)	1.09 (0.99)	0.047	0.27 (0.60)	3.44 (0.84)
AP	0.040	2.57 (0.11)	5.67 (0.58)	0.036	5.99 (0.01)	7.55 (0.37)	0.046	0.43 (0.51)	2.63 (0.92)	0.045	0.62 (0.43)	6.20 (0.52)
IG	0.044	1.12 (0.29)	3.40 (0.85)	0.035	6.71 (0.01)	7.68 (0.36)	0.048	0.15 (0.70)	3.88 (0.79)	0.046	0.43 (0.51)	3.58 (0.83)
FI	0.047	0.27 (0.60)	8.08 (0.33)	0.037	5.31 (0.02)	6.54 (0.48)	0.048	0.07 (0.79)	3.98 (0.78)	0.047	0.27 (0.60)	1.73 (0.97)
HY	0.045	0.62 (0.43)	5.78 (0.57)	0.037	4.67 (0.03)	6.96 (0.43)	0.047	0.27 (0.60)	3.50 (0.84)	0.049	0.02 (0.90)	3.04 (0.88)
	Turkey											
RM		0.41 (0.52)	15.89 (0.03)									
G	0.029	13.17 (0.00)	23.89 (0.00)									
E	0.045	0.62 (0.43)	5.90 (0.55)									
AP	0.029	13.18 (0.00)	20.04 (0.01)									
IG	0.03	12.10 (0.00)	15.55 (0.03)									
FI	0.037	4.67 (0.03)	11.36 (0.12)									
HY	0.043	1.42 (0.23)	5.70 (0.57)									

Table 5.12.a Results for 5% VaR Failure Rates – Asia

	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test
Model	Australia			China			Hong Kong		
RM	0.049	0.02 (0.90)	5.12 (0.65)	0.048	0.07 (0.79)	5.25 (0.63)	0.068	7.97 (0.00)	20.20 (0.01)
G	0.039	3.53 (0.06)	14.26 (0.05)	0.031	11.07 (0.00)	15.18 (0.03)	0.048	0.15 (0.70)	11.36 (0.12)
E	0.045	0.62 (0.43)	5.85 (0.56)	0.025	19.44 (0.00)	18.37 (0.01)	0.044	0.85 (0.36)	16.63 (0.02)
AP	0.044	1.12 (0.29)	5.14 (0.64)	0.044	0.85 (0.36)	13.25 (0.07)	0.043	1.42 (0.23)	16.42 (0.02)
IG	0.038	4.08 (0.04)	12.69 (0.08)	0.031	11.07 (0.00)	13.54 (0.06)	0.049	0.02 (0.90)	13.30 (0.07)
FI	0.042	1.76 (0.18)	11.92 (0.10)	0.044	0.85 (0.36)	8.18 (0.32)	0.057	1.30 (0.25)	12.30 (0.09)
HY	0.044	1.12 (0.29)	11.45 (0.12)	0.042	1.76 (0.18)	10.25 (0.17)	0.058	1.59 (0.21)	12.33 (0.09)
	India			Indonesia			Korea		
RM	0.056	0.79 (0.37)	14.71 (0.04)	0.042	1.76 (0.18)	16.15 (0.02)	0.065	5.53 (0.02)	20.12 (0.01)
G	0.039	3.53 (0.06)	11.73 (0.11)	0.037	5.31 (0.02)	25.51 (0.00)	0.052	0.15 (0.70)	10.74 (0.15)
E	0.040	3.03 (0.08)	13.62 (0.06)	0.034	7.49 (0.01)	24.80 (0.00)	0.055	0.58 (0.44)	23.38 (0.00)
AP	0.039	3.53 (0.06)	11.95 (0.10)	0.036	5.99 (0.01)	26.87 (0.00)	0.52	0.07 (0.80)	6.82 (0.45)
IG	0.038	4.08 (0.04)	13.08 (0.07)	0.038	4.08 (0.04)	21.68 (0.00)	0.053	0.26 (0.61)	7.94 (0.34)
FI	0.046	0.43 (0.51)	8.26 (0.31)	0.046	0.43 (0.51)	13.29 (0.07)	0.057	1.30 (0.25)	7.91 (0.34)
HY	0.046	0.43 (0.51)	8.27 (0.31)	0.042	1.76 (0.18)	18.17 (0.01)	0.055	0.58 (0.44)	6.95 (0.43)

Table 5.12.b Results for 5% VaR Failure Rates – Asia

	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test	Failure Rate	Kupiec Test	DQ Test
Model	Malaysia			Philippines			Singapore		
RM	0.049	0.02 (0.90)	4.15 (0.76)	0.053	0.26 (0.61)	5.74 (0.57)	0.056	1.03 (0.31)	6.41 (0.49)
G	0.029	13.18 (0.00)	17.12 (0.02)	0.033	9.18 (0.00)	11.16 (0.13)	0.042	1.76 (0.18)	5.60 (0.71)
E	0.035	6.71 (0.01)	15.25 (0.03)	0.034	7.49 (0.01)	14.53 (0.04)	0.043	1.42 (0.23)	6.41 (0.49)
AP	0.029	14.31 (0.00)	18.57 (0.01)	0.033	9.18 (0.00)	10.10 (0.18)	0.040	2.57 (0.11)	6.07 (0.53)
IG	0.029	13.18 (0.00)	17.11 (0.02)	0.033	9.18 (0.00)	11.09 (0.13)	0.040	3.03 (0.08)	5.45 (0.61)
FI	0.036	5.99 (0.01)	14.48 (0.04)	0.038	4.08 (0.04)	7.56 (0.37)	0.045	0.62 (0.43)	3.48 (0.84)
HY	0.039	3.53 (0.06)	10.44 (0.16)	0.039	3.53 (0.06)	6.79 (0.45)	0.044	0.85 (0.36)	3.08 (0.88)
	Taiwan			Thailand					
RM	0.056	0.79 (0.39)	11.52 (0.12)	0.054	0.41 (0.52)	15.32 (0.03)			
G	0.033	9.18 (0.00)	13.74 (0.06)	0.035	6.71 (0.01)	10.67 (0.15)			
E	0.034	7.49 (0.01)	18.67 (0.01)	0.037	5.31 (0.02)	11.12 (0.13)			
AP	0.031	11.07 (0.00)	14.85 (0.04)	0.033	8.31 (0.00)	12.77 (0.08)			
IG	0.037	4.67 (0.03)	8.14 (0.32)	0.038	4.08 (0.04)	8.09 (0.32)			
FI	0.039	3.53 (0.06)	13.51 (0.06)	0.042	1.76 (0.18)	16.16 (0.02)			
HY	0.037	4.67 (0.03)	17.75 (0.01)	0.041	2.14 (0.14)	14.24 (0.05)			

In summary table 5.13 a 5% VaR framework is used for the G7 countries. The RiskMetrics model again has the highest number of Failure rate cases compared to the GARCH type models. The average failure rate for the RiskMetrics model is 0.057 compared to the 0.046-0.049 for the GARCH type models. Looking at the significance of the Kupiec and Dynamic Quantile tests the RiskMetrics models does as well as the other models (none of the models return significant Kupiec statistics) whereas the significance of the DQ test is found for the RM only for 2 counties (Japan and USA) and no significance is found for the rest of the models with the exception of IGARCH model where in the case of Canada a significant DQ test statistic is observed.

Table 5.13 Summary of 5% VaR failure tests - G7

Model	Average Failure Rate	Significant Kupiec Test	Significant DQ test
G7			
RiskMetrics	0.057	None	Japan, USA
GARCH	0.048	None	None
EGARCH	0.049	None	None
APARCH	0.048	None	None
IGRACH	0.046	None	Canada
FIGARCH	0.047	None	None
HYGARCH	0.048	None	None

The summary table 5.14 examines the 5% VaR Failure rates within a sample of European markets. No surprises are reported for the Failure rate of the RiskMetrics which is above all other models, with an average rate of 0.055 contrasted to the average failure rate range of 0.0441 and 0.044 for the GARCH type models. A mixed picture is given in the next set of comparisons, where based on the Kupiec test RiskMetrics jointly with HYGARCH appear to be the best performers but based on the DQ test for several markets of the sample a significant test statistic is reported. On the other hand the HYGARCH is favoured by the two test statistics.

Table 5.14 Summary of 5% VaR failure tests - Europe			
Model	Average Failure Rate	Significant Kupiec Test	Significant DQ test
Europe			
RiskMetrics	0.055	Sweden, Switzerland	Austria, Denmark, Greece, Netherlands, Sweden, Switzerland, Turkey
GARCH	0.042	Greece, Spain, Turkey	Greece, Netherlands, Turkey
EGARCH	0.043	Greece, Netherlands Spain	Greece, Iceland, Netherlands
APARCH	0.041	Greece, Spain, Turkey	Greece, Netherlands, Turkey
IGRACH	0.042	Belgium, Greece, Netherlands, Spain, Turkey	Greece, Netherlands, Turkey
FIGARCH	0.043	Belgium, Finland, Spain, Turkey	Greece, Iceland, Netherlands
HYGARCH	0.044	Finland, Spain	Greece, Iceland

Summary table 5.15 shows a similar trend to that described above but for the Asian markets. The highest Failure rate, with an average of 0.054 compared to the average range for the GARCH models of 0.038-0.044 and similarly mixed as in table 5.14 overall picture when the Kupiec and DQ tests are considered.

Table 5.15 Summary of 5% VaR failure tests - Asia			
Model	Average Failure Rate	Significant Kupiec Test	Significant DQ test
Asia			
RiskMetrics	0.054	Hong Kong, Korea	Hong Kong, India, Indonesia, Korea, Thailand
GARCH	0.038	China, Indonesia, Malaysia, Philippines, Taiwan, Thailand	China, Indonesia, Malaysia
EGARCH	0.039	China, Indonesia, Malaysia, Philippines, Taiwan, Thailand	China, Hong Kong, Indonesia, Korea, Malaysia, Philippines, Taiwan
APARCH	0.039	Indonesia, Malaysia, Philippines, Taiwan, Thailand	Hong Kong, Indonesia, Malaysia, Taiwan
IGRACH	0.039	Australia, China, India, Indonesia, Malaysia, Philippines, Taiwan, Thailand	Indonesia, Malaysia
FIGARCH	0.045	Malaysia, Philippines	Malaysia, Thailand
HYGARCH	0.044	Taiwan	Indonesia, Taiwan

Again recalling that the main aim of this chapter is to assess the usefulness of the RiskMetrics model against the GARCH models, taking an overview of both the 1% and 5% VaR results one striking point emerges. For the 1% VaR the RM model does a particularly poor job. In fact for all series across the full range of markets, with the exception of China and Iceland the RM model performs the worst in terms of VaR exception frequency and with respect to both the Kupiec and DQ tests. However, at the 5% VaR level then the RM models performs as well as the best GARCH model in terms of the insignificance of the Kupiec and DQ tests.

The 5% VaR results suggest that across all markets the majority of the models estimated produce adequate VaR forecasts with insignificant Kupiec and DQ tests, with only a few exceptions.

5.5 Summary and Conclusion

The aim of this chapter is to assess whether the RiskMetrics volatility model can provide adequate forecasts of volatility in a Value-at-Risk setting in comparison to GARCH models. Academic research has highlighted the inherent flaws within the RiskMetrics model and in particular with respect to the undefined unconditional variance and the model's inability to produce long-horizon forecasts. The same research has highlighted that the GARCH model, which does not suffer from these drawbacks, is better able to capture the inherent time-dependency within volatility. However, the important question from a practitioner's point of view is whether this in-sample superiority of the GARCH model carries over to out-of-sample forecasting, or whether forecasts from the RiskMetrics model, which are easier to construct, perform as well as those of the GARCH model.

Using a selection of thirty-one international stock markets including those of the G7, thirteen further European markets and eleven further Asian markets RiskMetrics forecasts were compared to those of the GARCH model within a VaR framework. A simple conclusion is reached. In calculating 1% VaR then the APARCH model is preferred, while in calculating the 5% VaR the RiskMetrics model is adequate. The findings suggest that the RiskMetrics model is adequate in providing volatility forecasts when calculating the 5% Value-at-Risk for all markets. However, the

APARCH model is superior in obtaining the 1% Value-at-Risk forecasts.

Assessing the usefulness of the RiskMetrics model against the GARCH models, taking an overview of both the 1% and 5% VaR results a striking point emerges. For the 1% VaR the RiskMetrics model does a particularly poor job in contrast to the 5% VaR level where it performs as well as the best GARCH model in terms of the significance of the Kupiec and DQ tests. Specifically, for the 1% VaR for all the series across the full range of markets, with the exception of China and Iceland the RiskMetrics model performs the worst in terms of VaR exception frequency and with respect to both the Kupiec and DQ tests.

Looking at the results in more detail it is evident that the 1% VaR for the G7 the RiskMetrics model performs poorly. Its average failure rate is higher than the 1% VaR level for all markets. It also suffers from autocorrelated VaR exceptions. For the GARCH genre, all models achieve a lower average failure rate than the RiskMetrics models and indeed have similar values to each other. In terms of the specification tests both the APARCH and HYGARCH model perform the best, with only two markets significant on the Kupiec test and three markets significant on the DQ test. Examining the European markets, similar results occur, that is, the RiskMetrics model performs poorly, having the highest average failure rate across the thirteen European markets and having significant Kupiec and DQ tests. For the GARCH group of models, again the APARCH model performs well, having the lowest average failure rate. Moreover, it performs better than alternate GARCH models on the basis of the Kupiec test (joint with the IGARCH model) and the DQ test. Finally, the results for the Asian markets are again comparable with the G7 and also European results. The

RiskMetrics has the highest average failure rate, while it also results in significant specification tests for the majority of the markets. Although, it should be noted that not all markets return significant test statistics and on the basis of the DQ test only, then the RiskMetrics model performs favourably compared with the FIGARCH and HYGARCH models. With respect to the GARCH set of models, notably the GARCH, APARCH and IGARCH models perform the best across the average failure rate of the Kupiec and DQ tests.

Examining the 5% VaR results, it can be seen that the majority of the models perform well. Notably, the performance of the RiskMetrcs model is substantially improved upon the 1% VaR results. In particular, for the G7 markets, although the average failure rate remains higher for the RiskMetrics model, its performance based upon the Kupiec and DQ tests is only marginally inferior across these markets. That is, whereas as several of the GARCH models return insignificant specification tests for both tests, the RiskMerics model has a significant DQ test for two markets. With the exception of the IGARCH model, the alternate models within the GARCH class of model perform at a similar level. For the European and Asian markets, it can be argued that the RiskMetrics models perform as well, if not better, than the majority of the GARCH models. Notably, for both the European and Asian markets the GARCH model returns a VaR exception failure rate below the specified rate of 5%, and often significantly so. On the basis of the proximity of the failure rate to 5% and the Kupiec and DQ tests, the preferred models for both these groundings of markets is the RiskMetrics, FIGARCH and HYGARCH models.

6. A volatility forecasting exercise with VIX and Volume[45]

Abstract

The effect of two further parameters are considered in this thesis on improving volatility forecasts: the effect of the Volatility Index (VIX) and Trading Volume - volume in terms of trading value or in terms of trading quantities, on volatility forecasting. Including VIX and Volume components as exogenous variables within the setting of a selection of GARCH type models we discover that both VIX and Volume do a good job in improving our forecasts, however when VIX and Volume are considered together the results are improved further. In answering the question whether VIX produces better forecasts than the GARCH type models, the answer is no but the informational content of VIX cannot be ignored.

6.1 Introduction

In this chapter we wish to explore the effect of the Volatility Index (VIX) and Trading Volume on volatility forecasting. Both VIX and Volume have individually been considered in forecasting exercises in a large number of studies, with the aim of improving volatility forecasts. However, only a very small number of relatively recent

[45] This chapter was presented as a working paper at the BAFA Scottish Area Meeting at the University of Edinburgh Business School on the 31st of August 2011 and at the BAFA Doctoral Colloquium at Aston Business School on the 11th-12th of April 2011.

studies have investigated the impact of both factors together when trying to improve on the forecasting process.

Intuitively it is reasonable to consider VIX in such exercises because VIX is a benchmark of expected short-term market volatility, it is a forward–looking measure of volatility providing a benchmark upon which futures and options contracts on volatility can be written. VIX carries benefits for both practitioners and academics alike; first as an updated proxy for the future stock market volatility it is of high value for day-to-day trading decisions and second gives a better insight into risk and return patterns. Furthermore, because VIX contains market expectations it has been proven to be a useful instrument for forecasting volatility.

On the other hand Trading Volume also appears to have some interesting and useful properties which could help improve the accuracy of volatility forecasts. Volume in terms of trading value or trading quantities is caused by information flow. Two main theories exist looking at how this information flow is received by the market but what is commonly found is that a positive relationship between prices and volume exists hence a relationship between Trading Volume and volatility could exist.

The topic of volatility forecasting is not new within the finance literature and several aspects for producing accurate forecasts have been looked into. The focus of the literature has mainly been on the type of models used in order to produce accurate volatility forecasts capturing the different features found in the datasets used. Due to this, a significant number of different models have been proposed and comparisons have been made, see Chapters 2 & 3. Another key parameter for producing volatility

forecasts is the size of the in-sample period used for the forecasts. This was the topic of chapter 4 where the question of how much previous data do we need in order to improve volatility forecasts was investigated. With the use of recursive forecasts, for the first time within the volatility forecasting literature, some light was shed on the differing views between practitioners and academics, establishing that large in-sample periods do not necessarily improve the forecast ability of the models used, supporting the views of the practitioners. The improvement of volatility forecasts was also explored within a risk management setting, more specifically within a VaR setting where comparisons between the RiskMetrics approach and the 'more academic models' belonging to the GARCH genre were made in Chapter 5. The findings suggested that the GARCH type models provide more accurate results.

Building on from the previous chapters, here VIX and Volume data are used within a GARCH type model environment in order to establish whether there is added value when incorporating these two new datasets within the forecasting process. Three main markets are selected the UK, France and the USA, mainly due to data availability. The results suggest that both VIX and Volume improve on the informational content of the GARCH type models and more specifically it is proven that VIX does a better job in this process than Volume, moreover the results are further improved when both VIX and Volume are used together. However the trade off is between the statistical and economic significance of the findings since on one hand statistically the results are improved, on the other economically the value is minimal. In answering the question whether VIX produces better forecasts than the GARCH genre of models, the answer is no but the informational content of VIX cannot be ignored. The rest of this chapter is structured as follows; sections 2 and 3 provide information on VIX and

Volume respectively, followed by section 4, the data and methodology section, in section 5 the empirical results are presented with some concluding remarks are made in section 6.

6.2 Volatility Index VIX

The options market is a good source of information about volatility, Engle (2003). Comparing the volatility of VIX to the volatility of GARCH it becomes apparent that although the pattern is similar VIX is higher than GARCH (see the figure 6.1 below) and this could be because of two reasons. First, the option pricing relation might not be correct and does not allow for volatility risk premia on non-normal returns leading to higher option prices. Second, basic GARCH models have very limited information sets and do not use information on earnings or other latest information; hence the volatility forecasts by traders should be generally superior, Engle (2003).

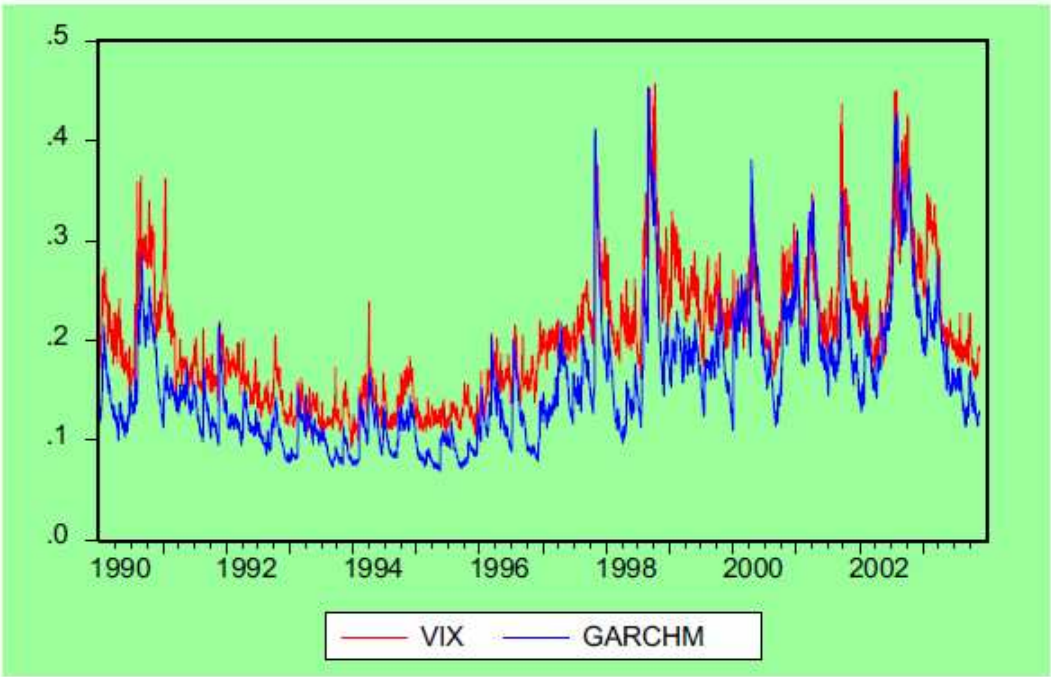


Figure 6.1. Source Engle (2003)

Simons (2003), describes the VIX as the "fear index" because its level indicates how much market participants are willing to pay in terms of implied volatility to hedge stock portfolios with S&P 100 index put options or to be long by buying S&P 100 index call options. In addition extreme values of VIX are seen as trading signals, for example with very high levels of VIX indicating that markets are pessimistic whereas a very low VIX leading to an increase in stock prices.

The original VIX was based on the Chicago Board Options Exchange (CBOE) Market Volatility Index and calculated as an average of the S&P 100 option implied volatilities and is computed on a real-time basis during the trading day which measures volatility instead of price, Fleming et al. (1995). According to Whaley (2009) VIX was introduced in 1993 for two reasons. First to provide a benchmark of expected short-term market volatility and second to provide an index upon which futures and options contracts on volatility could be written. When trying to understand VIX it is important to recognise that it is a forward–looking measure of volatility that investors expect to see.[46] Over the years due to a number of factors the calculation of the VIX has changed, now it is based on the S&P 500 index because it is a better known index and also because futures contracts on the S&P 500 are actively traded and S&P 500 option contracts are European-style making them easier to value, hence the VIX is implied by the current prices of the S&P 500 index options and represents expected future market volatility over the next 30 calendar days (Whaley, 2009).

An early study looking into how to improve forecasts using an (volatility) index was by LeBaron (1992) who used the volatility index but also suggested that other indices

[46] Whaley (2009) explains that VIX should be seen the same way as a bond's yield to maturity.

could be considered. However he does recognise and stresses the importance of searching for an 'optimal' volatility index which would help improve volatility forecasts.

The accuracy of volatility forecasts has been the topic of extensive research (see previous chapters) and as an alternative to GARCH volatility forecasts several academics have proposed the use of implied volatilities from options. The forecast ability of VIX has been explored in several studies concluding in most cases that the VIX index forecasts future volatility better than any historical volatility measure Ahomiemi (2008).

Fleming et al. (1995) find that the VIX performs better in forecasting future volatility than other historical measures. In addition, the importance of VIX was highlighted with benefits for both practitioners and academics. More specifically they mention that as an updated proxy for the future stock market volatility it is of high value for day-to-day trading decisions such as asset allocation, portfolio and risk management; conversely academics have the opportunity for a better insight into risk and return patterns. Furthermore, because VIX contains market expectations it was proven to be a useful instrument for forecasting volatility.

Blair et al. (2001) also reached the same conclusion that all relevant information is provided by the VIX index and that the VIX index provides the most accurate forecasts for all forecast horizons and for all performance measures used. As they mention (Blair et al. 2001) it is reasonable to compare the forecasting ability of

GARCH with implied volatility which are known to covary with realised volatility.[47] However, while Blair et al. (2001), Christensen and Prabhala (1998), Fleming (1998), Fleming et al. (1995), Hol and Koopman (2002) and Szakmary et al. (2002) find that option implied volatilities dominate over time series forecasts, other research provides contrary results in support of VIX. Corrado and Miller (2005) concluded that although VIX yields upward biased forecasts there are still more accurate than those of other historical models and Dennis et al. (2006) find that daily VIX changes are significant in predicting future index return volatility. Carr and Wu (2006) find that the VIX can predict movements in future realised variance, and that GARCH volatilities do not provide extra information once the VIX is included as a regressor. Giot and Laurent (2007), find that implied volatility has very high information content, even when extended decompositions of past realised volatility are used, this is also confirmed when adding GARCH-type volatility forecasts in the regressions. In contrast the results by Becker et al. (2006) state a differing view to the above mentioned findings and report that VIX is not an efficient volatility forecast and that other information can improve upon the VIX as a volatility forecast. Studies within a GARCH setting also produced mixed results with Day and Lewis (1992) finding that implied volatilities perform well but not better than the GARCH forecasts, but also that combinations of the two outperform univariate forecasts. In addition studies by Ederington and Guan (1999, 2002) and Martin and Zein (2002) find that GARCH models and historical volatility models do a good job, while Canina and Figlewski (1993) find that implied volatilities provide poor forecasts and that simple historical models perform better. In a related line of research poor forecast performance for GARCH type models have been reported extensively in the literature see for example

[47] Latane and Rendleman (1976) and Chiras and Manaster (1978).

Akgiray (1989), Boudoukh et al., (1997), Brailsford and Faff (1996), Dimson and Marsh (1990), Frennberg and Hansson (1996), Figlewski (1997), Heynen and Kat (1994), Jorion (1995), Schwert (1989, 1990a), and Schwert and Seguin (1990). This research often reports very low R^2's, in most cases less than 10% in a regression of true volatility on forecast volatility. Andersen and Bollerslev (1998) address the problem of low R^2's, they prove that regression methods will give low R^2 values when daily squared returns measure true volatility, even for optimal GARCH forecasts, because squared returns are noisy estimates of volatility. They show that intraday returns can be used to construct a realised volatility series that essentially eliminates the noise in measurements of daily volatility. They find remarkable improvements in the forecasting performance of GARCH models for Foreign Exchange data when they are used to forecast the realised series, compatible with theoretical analysis.

As can be seen the literature has yet to agree on the usefulness of VIX when forecasting. The purpose of this chapter is also to address this issue and perform further volatility forecasting exercises assessing the usefulness of VIX within a GARCH framework, as originally done by Blair et al. (2001). In the next section a further parameter, Trading Volume is considered.

6.3 Volume (VO and VA)

The relationship between stock market volume and volatility has been a subject of interest in the finance literature for several decades now. Karpoff (1987) produced a

systematic survey of this relationship reviewing a large number of previous studies using a variety of data sets and data frequencies from different markets arriving to the conclusion that "*volume is positively related to the magnitude of the price change and, in equity markets, to the price change per se*" (p. 109) hence there exists a positive contemporaneous correlation between the absolute price and volume measures. Since then several papers have been written on the topic.

A number of studies were also conducted and theoretical models of volume and volatility were proposed. Following from the work of Clark (1973) the 'Mixture of Distributions Hypothesis' (MDH) suggests that the volume and volatility should be positively correlated since both originate from the same source; the rate of information flow. Representative studies of this hypothesis are Epps and Epps (1975), Tauchen and Pitts (1983), Harris (1986, 1987), and Andersen (1996). The MDH implies that past volume does not contain any additional useful information on the future dynamics of volatility.

On the other hand a different class of models known as 'Sequential Information Hypothesis' (SIH) originated from the work of Copeland (1976) and others, Jennings et al. (1981) and Smirlock and Starks (1984). This class of models advocates that new information enters the market sequentially implying that a bidirectional causality or positive contemporaneous relationship between volume and volatility could exist. In the same line of thought the Noise Trading Hypothesis (Milton and Raviv, 1993, Brock and LeBaron, 1996) suggests that a causal relationship exists which can be exploited for forecasting purposes. More recently Abu Hassan Shaari Mohd and Chin Wen (2007) study the dynamic relationships of the realised volatility and trading

volume using a bivariate vector autoregressive methodology. Their empirical results support the MDH, however they also discovered significant causal relations between trading volume and return volatility in accordance with the SIH. Nevertheless, it has been argued by Wang (1994), that information asymmetry and investor heterogeneity could also be a factor in the above relationship.[48]

According to Taylor (2008) a number of studies have demonstrated that the performance of volatility models can be significantly improved with the inclusion of proxies of information flow in their model specification. Trading volume is one of those factors which have shown to lead to significant improvements (Karpoff, 1987; Lamoureux and Lastrapes,1990; Bessembinder and Seguin, 1993; Bollerslev and Jubinski, 1999; Luu and Martens, 2003).

The main focus of the majority of studies conducted was aimed at assessing trading volume as an information proxy in relation to the volatility in prices (Heimstra and Jones, 1994; Lamoureux and Lastrapes, 1994; Richardson and Smith, 1994). A limited number of studies investigated the information content of trading volume in volatility forecasting applications. Initial results reported on this were discouraging concluding that trading volume was not helpful in improving forecasts and that it cannot forecast volatility directly Lamoureux and Lastrapes (1994), and Brooks (1998). It was demonstrated that trading volume may not be the most accurate measure of information flow because volume could be liquidity motivated or occur as a result of divergent in trader opinion Taylor (2008).

[48] Volume absolute returns relationship, volume is positively correlated with absolute returns and this correlation is increased by information asymmetry.

A relatively different story emerges with some initial signs of success when trading volume is used within a GARCH type setting. Lamoureux and Lastrapes (1990) use daily trading volume as a proxy for information arrival within a GARCH (1,1) framework and find that daily transactions volume has a significant explanatory power on the variance of daily returns and that ARCH effects tend to disappear when volume is included in the variance equation. Wagner and Marsh (2005) successfully managed to explain the heteroskedasticity in returns using trading volume by extending the work of Lamoureux and Lastrapes (1990) and adopting an asymmetric GARCH in-mean model specification of Golsten et al. (1993).

Brooks (1998) focussed onto the causal relationship between volatility and trading volume with the use of a number of different statistical models (and within an GARCH type setting) on the New York Stock Exchange market. He found that lagged stock market volume measures play a little role in improving the out-of-sample forecasting performance of volatility models.

In a more recent study by Donaldson and Kamstra (2005) who also use a GARCH framework renowned for its special attribute for capturing volatility persistence,[49] it was found that lagged volume has no marginal power to forecast future volatility; more specifically: "*results from pervious research suggests that in an ARCH model that already accounts for the impact of lagged return innovations on future volatility, lagged volume will have no marginal power to forecast future volatility*" (p.1). Donaldson and Kamstra (2005) use a forecast combination approach in the same

[49] Shocks to the conditional variance showed a high degree of persistence.

study adopting lagged volume with option implied volatility within a GARCH framework and find that trading volume is significant.

There were concerns raised with regard to the inclusion of volume acknowledged by Lamoureux and Lastrapes (1990) and by Brooks (1998) in addressing the possible problem of simultaneity bias. For this reason a lagged volume component was used. An additional concern is also raised within the GARCH framework for which the same technique proved to be the answer (a lagged volume component). The expected variance of returns of the GARCH type models is generated as a polynomial of past squared returns, this however imposes a weakness in the case where the previous period return was zero, the lagged square return would also be zero without taking into account any price fluctuations between the two periods, Fuertes et al. (2008). As mentioned above, a way forward is to augment the GARCH type model with such variables that carry predictive power for future volatility; lagged volume is an appropriate instrument for contemporaneous volume. This is the methodology that will be followed in the next sections.

6.4 Data and Methodology

6.4.1 Data

Finding data for this exercise was not an easy task. The main constraint was finding VIX data for the indices used in the previous chapters. According to Whaley (2009) the CBOE methodology for computing the VIX is not unique to the prices of S&P

500 options but can be applied to any index option market. Examples of indices who have used the CBOE methodology are for the NASDAQ 100 the “VXN” and for the DJIA the “VXD”. NYSE Euronext has applied the same methodology to other option indices on European indices for example, on the AEX (an index of 25 stocks traded in Amsterdam), BEL20 (an index of 20 stocks), the CAC40 (an index of 40 French stocks) and the FTSE 100 (an index of 100 stocks traded in the United Kingdom).

For the purposes of this chapter three representative indices are obtained restricted on the amount of available data. Data from the USA, France and the UK were obtained from Datastream allowing a sufficient amount of observations. Due to data availability for the USA, the data ranges from the beginning of 1990 till the end of 2010, for France and the UK the data covers only the first decade of 2000 (from the beginning of 2000 till end of 2010). For each country the relevant index (daily closing prices), VIX data and trading volume data were obtained. The measure of trading volume both volume in terms of number of trades and transaction value were considered for the reason of data availability, more specifically for the same sample period for the USA and France volume in terms of traded quantities is used (VO) and for the UK volume in terms of transaction value is considered (VA).

The first step is to convert the prices into logarithmic returns and report the descriptive statistics. Descriptive statistics for VIX and Volume are also presented.

Table 6.1 Descriptive statistics for returns FTSE, CAC and S&P			
	RFTSE	RCAC	RS_P
Mean	-6.39E-05	-0.000150	0.000224
Median	0.000000	0.000000	0.000187
Maximum	0.093843	0.105946	0.109572
Minimum	-0.092656	-0.094715	-0.094695
Std. Dev.	0.013137	0.015561	0.011529
Skewness	-0.123957	0.063716	-0.198720
Kurtosis	9.259858	8.190561	12.24104
Jarque-Bera	4570.674	3184.433	19460.14

The markets appear to be similar with very small differences for the reported statistics. Excess positive kurtosis is present for all countries indicating thicker tails from the normal distribution.

Table 6.2 Descriptive statistics for VIX and Volume						
	UK		France		USA	
	VIX	Volume	VIX	Volume	VIX	Volume
Mean	21.91006	1457075	24.43638	4171692	20.39075	462.5611
Median	20.13750	1426356	22.67300	3795400	19.060000	496.9100
Maximum	75.54000	4461012	78.05000	16017700	80.86000	836.1900
Minimum	9.099000	93760	9.242000	164218	9.310000	163.8800
Std. Dev.	9.610590	472912.8	9.896830	1657460	8.236873	177.3889
Skewness	1.574084	0.623678	1.511701	1.482752	1.996063	-0.086096
Kurtosis	6.589943	4.786143	5.955547	7.623696	10.11884	1.910078
Jarque-Bera	2574.348	535.9256	2065.448	3486.214	15154.91	276.9997

The same can be said also for the VIX and Volume data of the sample countries, all statistics appear to be similar.

Table 6.3 Correlation Coefficient Matrix		
UK	VIX	Volume
VIX	1	-0.016407
Volume	-0.0156407	1
France	VIX	Volume
VIX	1	-0.03184
Volume	-0.03184	1
US	VIX	Volume
VIX	1	0.075148
Volume	0.075148	1

The correlation coefficient matrices are presented below and as can be observed the coefficients are small hence suggesting no correlation and hence multicollinearity will not be present.

For France and the UK the in-sample period is from the beginning of 2000 till the end of 2006 and for the USA from the beginning of 1990 till the end of 2006.

6.4.2 Methodology

The testing procedure of Mincer-Zarnowitz (MZ) is used where the true volatility value is regressed on the forecast value and the coefficient of determination is obtained R^2 for comparison purposes.

$$\sigma_t^2 = \alpha + \beta h_t^2 + \varepsilon_t \tag{6.1}$$

As a measure of true volatility the procedure of Pagan and Schwert (1990) is followed, where true volatility is used as a proxy by the squared error from the conditional mean model for returns estimated over the whole sample. The next step is twofold. In the first testing procedure lagged values of VIX and Volume are included only on the right hand side of the MZ regression as explanatory variables and the following regressions are run in addition to the regression 6.1 which will act the point of reference. In the second part of this procedure lagged values of VIX act as an exogenous variable in the variance equation and then the OLS regressions (6.1 – 6.4) are run obtaining the coefficient of determination R^2 which is reported for all the

regressions, representing the information content of the particular model used. A higher R^2 value is preferred.

$$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \varepsilon_t \quad (6.2)$$

$$\sigma_t^2 = \alpha + \beta h_t^2 + \delta Volume_{t-1} + \varepsilon_t \quad (6.3)$$

$$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \delta Volume_{t-1} + \varepsilon_t \quad (6.4)$$

A selection of six models is used each, belonging to the GARCH genre, capturing the main features found in empirical investigations of financial market returns such as volatility clustering, information asymmetry and long memory. The models considered include the GARCH model (the first generation of ARCH models, the Generalised ARCH symmetric model), the second generation asymmetric models in particular the TGARCH by Glosten, Jagannathan and Runkle (1993), the EGARCH by Nelson (1991), and the APARCH by Ding, Granger and Engle (1993) and the third generation long-memory models the CGARCH by Engle and Lee (1999) and the IGARCH by Engle and Bollerslev (1986).

Modelling

In this section a small reminder of the models to be used in this chapter are presented. These models have extensively been looked into in literature review section 2.6 (page 42) and in Chapter 3 section 3.3.2 (page 69).

The GARCH(1,1) model of Engle (1982) and Bolerslev (1986) is given by:

$$h_{t+1}^2 = \omega + \alpha \varepsilon_t^2 + \beta h_t^2 \quad (6.5)$$

The Threshold-GARCH (TGARCH) model of Glosten, Jagannathan and Runkle (1993), is given by:

$$h_{t+1}^2 = \omega + \alpha\varepsilon_t^2 + \gamma\varepsilon_t^2 I_t + \beta h_t^2 \tag{6.6}$$

The EGARCH model of Nelson (1991) is given by:

$$\log(h_{t+1}^2) = \omega + \alpha\left|\frac{\varepsilon_t}{h_t}\right| + \gamma\frac{\varepsilon_t}{h_t} + \beta\log(h_t^2) \tag{6.7}$$

The component GARCH (CGARCH) model of Engle and Lee (1999) is specified as:

$$h_{t+1}^2 = q_{t+1} + \alpha(\varepsilon_t^2 - q_t) + \beta(h_t^2 - q_t) \tag{6.8}$$

The asymmetric power-ARCH (APARCH) by Ding et al. (1993) is specified as:

$$h_t^\delta = \omega + \alpha_1(|\varepsilon_{t-1}| - \gamma\varepsilon_{t-1})^\delta + \beta_1 h_{t-1}^\delta \tag{6.9}$$

The IGARCH is specified as an extension to GARCH model for which the following condition must hold, $\alpha + \beta = 1$ for the conditional variance to be clearly non-stationary.

6.5 Empirical results

The results in the tables below (see column A of tables 6.3 – 6.5), present us with a general pattern irrespective or the type of model or the index used. The inclusion of a lagged VIX component alone increases the value of the R^2 statistic compared to the R^2 of the original MZ regression (equation 6.1). The same happens when a lagged Volume component is added to the original MZ regression, however the coefficient of determination is further increased when both a lagged VIX and lagged Volume component combined are included in the MZ regression. For the UK the VIX element adds to the informational content of all the forecast models by 0.1% to 1.6% and the Volume on the other hand adds to the accuracy by 0.05% to 1%. The combined VIX and Volume increases the R^2 by 0.5% to 3%. In the case of France the increases are between 0.05% and 3% when VIX is included in the MZ regression and 0.07% and 1.3% when Volume is included and when both combined between 0.7% and 6%. In the USA the increases in the R^2 are between 0.07% and 3.5%, 0.004% and 0.04%, and 0.026% and 5% when respectively, VIX, Volume and the two combined are added in the MZ regression.

The results of including lagged VIX component in the variance equation and then running, as before the same MZ regressions, are presented in column B of tables 6.3 and 6.5. The coefficients of determination change in both directions[50] for the different models with no consistent pattern emerging, leading to no consistent conclusion. For

[50] Two striking exceptions are present for the IGARCH model in the UK and France, where the coefficient of determination is significantly reduced. Further calculations were conducted for the construction of the IGARCH model including a lagged value of VIX in the variance equation. The number of iterations was increased from the default of 500 (to 1000, 1500 and finally 10000). This adjustment although improves and increases the value of the coefficient of determination the values are still relatively low. The original value of 0.000103 increases to 0.0014 and respectively the original value of 0.0151168 increases to 0.026904). The improved values are added in the tables. (0.006515 improves to 0.010310 and 0.037931 to 0.040185)

this reason and henceforth for this chapter a lagged VIX component will not be included in the variance equation of the GARCH type models.

The alleged poor performance of the GARCH type models has been the topic of several academic papers which criticised the very low –often below 10% R^2's, when as above the method of MZ regression analysis is used for measuring the accuracy of the forecasts; Akgiray (1989), Boudoukh et al. (1997), Brailsford and Faff (1996), Dimson and Marsh (1990), Frennberg and Hansson (1996), Figlewski (1997), Heynen and Kat (1994), Jorion (1995), Schwert (1989, 1990a), and Schwert and Seguin (1990). Although Andresen and Bollerslev (1998)[51] provide an explanation for the very low values of R^2 it was proven that the inclusion of a lagged VIX parameter in the MZ regression increases the value of the R^2 in the majority of the cases and by up to 3.5%, providing this way improved accuracy of the volatility forecasts confirming to some extent the work by Fleming et al. (1995), Blair et al. (1995) and Ahomiemi (2008) who argue that VIX index improves the accuracy of volatility forecasts when it is used as an instrument in the forecasting process.

Mixed results have also been reported with the use of Volume as a proxy of information flow within the volatility forecasting debate. It was demonstrated that lagged Volume improves the accuracy of forecasts produced as previously mentioned by Karpoff (1987), Lamoureux and Lastrapes (1990), Bessembinder and Seguin

[51] The often low R^2 's does not come as a surprise. Andersen and Bollerslev (1998) address the problem of low R^2's, by proving that regression methods will give low R^2 values when daily squared returns measure true volatility, even for optimal GARCH forecasts, because squared returns are noisy estimates of volatility. Used for comparison purposes, the low R^2's do not invalidate the findings, since the method adopted is generally accepted (Mincer and Zarnowitz, 1996). It would be desirable to use realised volatility measure but it is not feasible to obtain the necessary intra-day data for all the series.

(1993), Bollerslev and Jubinski (1999) and Luu and Martens (2003) and Taylor (2008). Our results confirm those of Brooks who found that lagged Volume within an GARCH setting played little role in improving the out of sample performance of the models and this becomes more apparent when the findings are compared to the improvements reported with the inclusion of a lagged VIX parameter. On the other hand the coefficients of determination are increased further when both lagged values of the VIX and Volume are included together in the MZ regressions.

Table 6.3 UK results		
Model and MZ regression/ Coefficient of determination	**A**	**B**
	R^2	R^2
GARCH(1,1)		
$\sigma_t^2 = \alpha + \beta h_t^2 + \varepsilon_t$	0.196733	0.200257
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \varepsilon_t$	0.213299	0.212880
$\sigma_t^2 = \alpha + \beta h_t^2 + \delta Volume_{t-1} + \varepsilon_t$	0.204900	0.206888
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \delta Volume_{t-1} + \varepsilon_t$	0.229070	0.227285
TGARCH		
$\sigma_t^2 = \alpha + \beta h_t^2 + \varepsilon_t$	0.250620	0.272672
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \varepsilon_t$	0.252049	0.273056
$\sigma_t^2 = \alpha + \beta h_t^2 + \delta Volume_{t-1} + \varepsilon_t$	0.256011	0.276214
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \delta Volume_{t-1} + \varepsilon_t$	0.259684	0.277637
EGARCH		
$\sigma_t^2 = \alpha + \beta h_t^2 + \varepsilon_t$	0.262685	0.223805
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \varepsilon_t$	0.262812	0.232449
$\sigma_t^2 = \alpha + \beta h_t^2 + \delta Volume_{t-1} + \varepsilon_t$	0.268218	0.232655
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \delta Volume_{t-1} + \varepsilon_t$	0.268340	0.245897
APARCH		
$\sigma_t^2 = \alpha + \beta h_t^2 + \varepsilon_t$	0.246791	0.261375
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \varepsilon_t$	0.247934	0.262488
$\sigma_t^2 = \alpha + \beta h_t^2 + \delta Volume_{t-1} + \varepsilon_t$	0.252534	0.267538
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \delta Volume_{t-1} + \varepsilon_t$	0.255843	0.267635
CGARCH		
$\sigma_t^2 = \alpha + \beta h_t^2 + \varepsilon_t$	0.199565	0.218239
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \varepsilon_t$	0.215767	0.226209
$\sigma_t^2 = \alpha + \beta h_t^2 + \delta Volume_{t-1} + \varepsilon_t$	0.207790	0.225483
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \delta Volume_{t-1} + \varepsilon_t$	0.231211	0.238858
IGARCH		
$\sigma_t^2 = \alpha + \beta h_t^2 + \varepsilon_t$	0.191066	0.000103
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \varepsilon_t$	0.208177	0.200822
$\sigma_t^2 = \alpha + \beta h_t^2 + \delta Volume_{t-1} + \varepsilon_t$	0.201705	0.015168
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \delta Volume_{t-1} + \varepsilon_t$	0.226230	0.218733

Notes:
Column A presents the results of the MZ regression
Column B presents the results of the MZ regression with VIX_{t-1} in variance equation

Table 6.4 France results		
Model and MZ regression/ Coefficient of determination	**A**	**B**
	R^2	R^2
GARCH(1,1)		
$\sigma_t^2 = \alpha + \beta h_t^2 + \varepsilon_t$	0.158798	0.165379
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \varepsilon_t$	0.189686	0.190220
$\sigma_t^2 = \alpha + \beta h_t^2 + \delta Volume_{t-1} + \varepsilon_t$	0.169663	0.175281
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \delta Volume_{t-1} + \varepsilon_t$	0.213198	0.213126
TGARCH		
$\sigma_t^2 = \alpha + \beta h_t^2 + \varepsilon_t$	0.211450	0.165379
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \varepsilon_t$	0.215672	0.190220
$\sigma_t^2 = \alpha + \beta h_t^2 + \delta Volume_{t-1} + \varepsilon_t$	0.219257	0.175281
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \delta Volume_{t-1} + \varepsilon_t$	0.228711	0.213126
EGARCH		
$\sigma_t^2 = \alpha + \beta h_t^2 + \varepsilon_t$	0.220740	0.208410
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \varepsilon_t$	0.221265	0.211627
$\sigma_t^2 = \alpha + \beta h_t^2 + \delta Volume_{t-1} + \varepsilon_t$	0.228986	0.221102
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \delta Volume_{t-1} + \varepsilon_t$	0.232145	0.228441
APARCH		
$\sigma_t^2 = \alpha + \beta h_t^2 + \varepsilon_t$	0.207078	0.211165
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \varepsilon_t$	0.211063	0.212135
$\sigma_t^2 = \alpha + \beta h_t^2 + \delta Volume_{t-1} + \varepsilon_t$	0.216454	0.223064
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \delta Volume_{t-1} + \varepsilon_t$	0.225680	0.227185
CGARCH		
$\sigma_t^2 = \alpha + \beta h_t^2 + \varepsilon_t$	0.160721	0.173206
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \varepsilon_t$	0.192003	0.194790
$\sigma_t^2 = \alpha + \beta h_t^2 + \delta Volume_{t-1} + \varepsilon_t$	0.172864	0.183289
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \delta Volume_{t-1} + \varepsilon_t$	0.214777	0.215493
IGARCH		
$\sigma_t^2 = \alpha + \beta h_t^2 + \varepsilon_t$	0.153394	0.006515
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \varepsilon_t$	0.186247	0.197821
$\sigma_t^2 = \alpha + \beta h_t^2 + \delta Volume_{t-1} + \varepsilon_t$	0.167076	0.037931
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \delta Volume_{t-1} + \varepsilon_t$	0.212849	0.214458

Notes:
Column A presents the results of the MZ regression
Column B presents the results of the MZ regression with VIX_{t-1} in variance equation

Table 6.5 USA results		
Model and MZ regression/ Coefficient of determination	**A**	**B**
	R^2	R^2
GARCH(1,1)		
$\sigma_t^2 = \alpha + \beta h_t^2 + \varepsilon_t$	0.216410	0.225906
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \varepsilon_t$	0.238215	0.235323
$\sigma_t^2 = \alpha + \beta h_t^2 + \delta Volume_{t-1} + \varepsilon_t$	0.216414	0.240116
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \delta Volume_{t-1} + \varepsilon_t$	0.251196	0.256492
TGARCH		
$\sigma_t^2 = \alpha + \beta h_t^2 + \varepsilon_t$	0.255646	0.272479
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \varepsilon_t$	0.260538	0.274240
$\sigma_t^2 = \alpha + \beta h_t^2 + \delta Volume_{t-1} + \varepsilon_t$	0.255655	0.274240
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \delta Volume_{t-1} + \varepsilon_t$	0.265013	0.277279
EGARCH		
$\sigma_t^2 = \alpha + \beta h_t^2 + \varepsilon_t$	0.271440	0.225906
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \varepsilon_t$	0.272148	0.235323
$\sigma_t^2 = \alpha + \beta h_t^2 + \delta Volume_{t-1} + \varepsilon_t$	0.271620	0.240116
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \delta Volume_{t-1} + \varepsilon_t$	0.274043	0.256492
APARCH		
$\sigma_t^2 = \alpha + \beta h_t^2 + \varepsilon_t$	0.258287	0.225906
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \varepsilon_t$	0.262248	0.235323
$\sigma_t^2 = \alpha + \beta h_t^2 + \delta Volume_{t-1} + \varepsilon_t$	0.258305	0.240116
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \delta Volume_{t-1} + \varepsilon_t$	0.266217	0.256492
CGARCH		
$\sigma_t^2 = \alpha + \beta h_t^2 + \varepsilon_t$	0.223984	0.235552
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \varepsilon_t$	0.242010	0.243087
$\sigma_t^2 = \alpha + \beta h_t^2 + \delta Volume_{t-1} + \varepsilon_t$	0.223993	0.237381
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \delta Volume_{t-1} + \varepsilon_t$	0.253374	0.254688
IGARCH		
$\sigma_t^2 = \alpha + \beta h_t^2 + \varepsilon_t$	0.197864	0.217967
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \varepsilon_t$	0.232348	0.238767
$\sigma_t^2 = \alpha + \beta h_t^2 + \delta Volume_{t-1} + \varepsilon_t$	0.198068	0.217973
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \delta Volume_{t-1} + \varepsilon_t$	0.249740	0.251438

Notes:
Column A presents the results of the MZ regression
Column B presents the results of the MZ regression with VIX_{t-1} in variance equation

6.6 Further analysis: A forecast encompassing exercise

The next step in the analysis is to examine the relative forecasting performance in order to identify whether the additional components carry additional information over the base forecast in this case the forecast produced by the GARCH type models. This method is known as forecast encompassing originally developed by Chong and Hendry (1986). Taking the MZ regressions previously produced, the sign and significance of the additional components (lagged VIX and lagged Volume) are assessed and inferences are made. The significance of the coefficients is established by the reported p-values of the lagged VIX coefficient (γ) and the lagged Volume coefficient (δ). For the results below all the GARCH type model forecasts acting as the base model are reported. Tables 6.6 – 6.8 show the UK coefficients for VIX and Volume and their respective p-values of the regressions reported above; refer to equations 6.1 – 6.4.

The results for the UK show that in all cases the β coefficients are greater in absolute terms than the γ coefficients and are all positive and significant in half of the cases. The coefficient γ is always positive with the exception of when the EGARCH model is used and not significant for the TGARCH, EGARCH and APARCH models, concluding that lagged VIX is not encompassed by half of the GARCH forecasts (GARCH, CGARCH and IGARCH) but some explanatory power is added, see table 6.6.

A different pattern emerges in the next set of MZ regressions, where the GARCH type forecast is regressed on the lagged Volume component. Table 6.6 presents all positive and all significant coefficients for both the GARCH type forecast β which also appear

to be of larger value than the lagged Volume δ coefficients which are very small. The results suggest that lagged Volume is not encompassed in the GARCH type forecasts.

Table 6.6 UK					
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \varepsilon_t$			$\sigma_t^2 = \alpha + \beta h_t^2 + \delta Volume_{t-1} + \varepsilon_t$		
Model	Coefficient	P-value	Model	Coefficient	P-value
GARCH			GARCH		
β	0.499887	0.0000	β	0.886197	0.0000
γ	1.48E-05	0.0000	δ	1.21E-10	0.0000
TGARCH			TGARCH		
β	0.902803	0.0000	β	1.011054	0.0000
γ	4.40E-06	0.1258	δ	9.62E-11	0.0049
EGARCH			EGARCH		
β	1.570538	0.0000	β	1.497417	0.0000
γ	-1.45E-06	0.6464	δ	9.51E-11	0.0051
APARCH			APARCH		
β	1.04573	0.0000	β	1.164020	0.0000
γ	4.09E-06	0.1720	δ	9.89E-11	0.0040
CGARCH			CGARCH		
β	0.517935	0.0000	β	0.886243	0.0000
γ	1.43E-05	0.0000	δ	1.21E-10	0.0000
IGARCH			IGARCH		
β	0.440251	0.0000	β	0.861586	0.0000
γ	1.59E-05	0.0000	δ	1.39E-10	0.0001

When both VIX and Volume are included in the MZ regression, their coefficients are positive and all significant. This suggests that jointly VIX and Volume carry additional information and are not encompassed by the GARCH type forecast.

Table 6.7 UK		
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \delta Volume_{t-1} + \varepsilon_t$		
Model	Coefficient	P-value
GARCH		
β	0.386332	0.0001
γ	1.86E-05	0.0000
δ	1.75E-10	0.0000
TGARCH		
β	0.804600	0.0000
γ	7.50E-06	0.0153
δ	1.22E-10	0.0006
EGARCH		
β	1.434049	0.0000
γ	1.54E-06	0.6550
δ	1.01E-10	0.0053
APARCH		
β	0.924316	0.0000
γ	7.41E-06	0.0217
δ	1.24E-10	0.0006
CGARCH		
β	0.412337	0.0000
γ	1.80E-05	0.0000
δ	1.73E-10	0.0000
IGARCH		
β	0.337167	0.0007
γ	1.96E-05	0.0000
δ	1.86E-10	0.0000

For France, as can be seen in the table below (table 6.8), all the γ coefficients for the lagged VIX component are positive and significant, with the exception of when the EGARCH (as in the case of the UK) is used, and all the β coefficients are all positive and significant except of the IGARCH model. The GARCH type forecasts do not encompass VIX. A similar pattern for France is followed when Volume is used in the MZ regressions. In all cases both VIX and Volume are not encompassed in the

GARCH type forecasts but carry additional information, also jointly as is shown bellow in table 6.9.

Table 6.8 FRANCE					
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \varepsilon_t$			$\sigma_t^2 = \alpha + \beta h_t^2 + \delta Volume_{t-1} + \varepsilon_t$		
Model	Coefficient	P-value	Model	Coefficient	P-value
GARCH			GARCH		
β	0.251535	0.0210	β	0.885790	0.0000
γ	2.69E-05	0.0000	δ	4.22E-11	0.0001
TGARCH			TGARCH		
β	0.805789	0.0000	β	1.056538	0.0000
γ	1.05E-05	0.0091	δ	3.50E-11	0.0010
EGARCH			EGARCH		
β	1.290984	0.0000	β	1.416506	0.0000
γ	4.15E-06	0.3557	δ	3.52E-11	0.0008
APARCH			APARCH		
β	0.901286	0.0000	β	1.206680	0.0000
γ	1.08E-05	0.0115	δ	3.38E-11	0.0003
CGARCH			CGARCH		
β	0.301209	0.0028	β	0.878353	0.0000
γ	2.55E-05	0.0000	δ	4.46E-11	0.0000
IGARCH			IGARCH		
β	0.155911	0.1615	β	0.856440	0.0000
γ	2.96E-05	0.0000	δ	4.76E-11	0.0000

Table 6.9 FRANCE		
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \delta Volume_{t-1} + \varepsilon_t$		
Model	Coefficient	P-value
GARCH		
β	0.085070	0.4488
γ	3.32E-05	0.0000
δ	6.53E-11	0.0000
TGARCH		
β	0.628405	0.0000
γ	1.67E-05	0.0001
δ	4.90E-11	0.0000
EGARCH		
β	1.035016	0.0000
γ	1.10E-05	0.0242
δ	4.45E-11	0.0001
APARCH		
β	0.686015	0.0000
γ	1.74E-05	0.0001
δ	5.16E-11	0.0000
CGARCH		
β	0.179444	0.0803
γ	3.04E-05	0.0000
δ	6.35E-11	0.0000
IGARCH		
β	0.018150	0.8728
γ	3.53E-05	0.0000
δ	6.70E-11	0.0000

And similar findings are reported also for the USA (Tables 6.10 and 6.11)

Table 6.10 USA					
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \varepsilon_t$			$\sigma_t^2 = \alpha + \beta h_t^2 + \delta Volume_{t-1} + \varepsilon_t$		
Model	Coefficient	P-value	Model	Coefficient	P-value
GARCH			GARCH		
β	0.373281	0.0010	β	0.986663	0.0000
γ	2.16E-05	0.0000	γ	1.34E-08	0.9375
TGARCH			TGARCH		
β	0.738032	0.0000	β	1.011436	0.0000
γ	1.02E-02	0.0037	γ	1.94E-08	0.9050
EGARCH			EGARCH		
β	1.34403	0.0000	β	1.524729	0.0000
γ	4.10E-06	0.2645	γ	9.07E-08	0.5741
APARCH			APARCH		
β	0.834736	0.0000	β	1.107517	0.0000
γ	9.24E-06	0.0088	γ	2.82E-08	0.8624
CGARCH			CGARCH		
β	0.456699	0.0000	β	0.996463	0.0000
γ	1.92E-05	0.0000	γ	2.08E-08	0.9021
IGARCH			IGARCH		
β	0.116864	0.3236	β	0.946135	0.0000
γ	2.87E-05	0.0000	γ	1.01E-07	0.5728

Table 6.11 USA		
$\sigma_t^2 = \alpha + \beta h_t^2 + \gamma VIX_{t-1} + \delta Volume_{t-1} + \varepsilon_t$		
Model	Coefficient	P-value
GARCH		
β	0.190245	0.1098
γ	3.42E-05	0.0000
δ	9.87E-07	0.0000
TGARCH		
β	0.597713	0.0000
γ	1.89E-05	0.0001
δ	6.08E-07	0.0053
EGARCH		
β	1.186478	0.0000
γ	1.04E-05	0.0388
δ	4.05E-07	0.0677
APARCH		
β	0.685624	0.0000
γ	1.77E-05	0.0002
δ	5.76E-07	0.0088
CGARCH		
β	0.289490	0.0122
γ	3.09E-05	0.0000
δ	9.24E-07	0.0000
IGARCH		
β	-0.031436	0.7937
γ	4.14E-05	0.0000
δ	1.11E-06	0.0000

Overall the findings suggest that VIX and Volume help us produce more accurate results than when applying GARCH type forecasts alone, and it was seen that GARCH forecasts do not encompass lagged values of VIX and Volume. The accuracy of the forecasts was measured using MZ regressions and by comparing the coefficients of determination. A higher coefficient of determination suggests that the informational content of the variables on the right hand side of the MZ regressions (lagged VIX and lagged volume) increases/improves the accuracy of the forecast, since the regressand (the variable on the left hand side) of the MZ regression is a forecast. The results indicated that the coefficient of determination increased, hence the informational content of the regressors helps improving the accuracy of the

forecasts. Furthermore when encompassing tests were performed it was found that the GARCH forecasts did not encompass the lagged values of VIX and Volume, suggesting that VIX and Volume could provide us with better forecasts than the GARCH models alone. The next section of this chapter addresses this issue by specifically looking at the forecast ability of VIX and comparing it to that of the GARCH models.[52] The question addressed next is, can VIX be used alone for forecasting and will VIX produce more accurate forecasts than the GARCH type models? To answer this question VIX is modelled as a basic AR(1) process of the form:

$$VIX_t = \alpha + \zeta VIX_{t-1} + \varepsilon_t \tag{6.10}$$

Then based on the VIX model forecasts are produced and further encompassing tests are performed. As a base model the GARCH type forecasts are used again, and the following MZ regression is run:

$$\sigma_t^2 = \alpha + \beta h_t^2 + \zeta VIX_t + \varepsilon_t \tag{6.11}$$

In the table below the results are reported in the same format as before. The coefficients β and ζ with their respective p-values and also for comparison purposes the R^2's are reported. For the UK, the results show positive and significant coefficients for the GARCH forecast and for the VIX also the coefficients are all positive with one exception (EGARCH) but not all are significant -in the cases where the TGARCH, EGARCH and APARCH are the base models. In half the cases the

[52] Volume is not used further because VIX produced better results in the previous exercise.

VIX forecast is not encompassed in the GARCH forecast, suggesting that in half the cases the VIX model has an explanatory power. Looking at the R^2 values, relatively high and as expected values are reported. However, what becomes apparent is the same pattern that was seen in the previous tables. The β coefficients not only are all significant but are also high, especially when compared to the ζ coefficient, and as before when compared to the γ and δ coefficients. This finding suggests that the GARCH type forecasts are adequate and although one cannot ignore the added value of including VIX -by increasing the R^2 values, the coefficients are so small and some times statistically insignificant that the superiority of the GARCH type models is confirmed. A similar story is concluded for France and the USA, where all the coefficients are positive implying that VIX carries additional information, however the coefficients are all very small.

Table 6.12

$$\sigma_t^2 = \alpha + \beta h_t^2 + \zeta VIX_t + \varepsilon_t$$

UK			FRANCE			USA		
GARCH								
	Coeff	P-value		Coeff	P-value		Coeff	P-value
α	-0.000237	0.0000	α	-0.000469	0.0000	α	-0.000359	0.0000
β	0.499887	0.0000	β	0.251535	0.0210	β	0.373281	0.0010
ζ	1.50E-05	0.0000	ζ	2.70E-05	0.0000	ζ	2.20E-05	0.0000
R^2	0.213299		R^2	0.189686		R^2	0.238215	
TGARCH								
	Coeff	P-value		Coeff	P-value		Coeff	P-value
α	-6.40E-05	0.2499	α	-0.000196	0.0127	α	-0.000174	0.0075
β	0.902803	0.0000	β	0.805789	0.0000	β	0.738032	0.0000
ζ	4.46E-06	0.1258	ζ	1.06E-05	0.0091	ζ	1.03E-05	0.0037
R^2	0.252049		R^2	0.215672		R^2	0.260538	
EGARCH								
	Coeff	P-value		Coeff	P-value		Coeff	P-value
α	-1.00E-05	0.8589	α	-0.000124	0.1303	α	-0.000106	0.1028
β	1.570538	0.0000	β	1.290984	0.0000	β	1.344030	0.0000
ζ	-1.47E-06	0.6464	ζ	4.19E-06	0.3557	ζ	4.18E-06	0.2645
R^2	0.262812		R^2	0.221265		R^2	0.272148	
APARCH								
	Coeff	P-value		Coeff	P-value		Coeff	P-value
α	-6.83E-05	0.2266	α	-0.000209	0.0102	α	0.000159	0.0152
β	1.045733	0.0000	β	0.901286	0.0000	β	0.834736	0.0000
ζ	4.14E-06	0.1720	ζ	1.09E-05	0.0115	ζ	9.40E-06	0.0088
R^2	0.247934		R^2	0.211063		R^2	0.262248	
CGARCH								
	Coeff	P-value		Coeff	P-value		Coeff	P-value
α	-0.000230	0.0000	α	-0.000446	0.0000	α	-0.000322	0.0000
β	0.517935	0.0000	β	0.301209	0.0028	β	0.456699	0.0000
ζ	1.45E-05	0.0000	ζ	2.57E-05	0.0000	ζ	1.96E-05	0.0000
R^2	0.215767		R^2	0.192003		R^2	0.242016	
IGARCH								
	Coeff	P-value		Coeff	P-value		Coeff	P-value
α	-0.000251	0.0000	α	-0.000512	0.0000	α	-0.000470	0.0000
β	0.440251	0.0000	β	0.155911	0.1615	β	0.116864	0.3236
ζ	1.61E-05	0.0000	ζ	2.99E-05	0.0000	ζ	2.92E-05	0.0000
R^2	0.208177		R^2	0.187530		R^2	0.232348	

As a final consideration a direct comparison between the coefficients of determination is performed for all the models by running the following two MZ regressions:

These are:

$$\sigma_t^2 = \alpha + \beta h_t^2 + \varepsilon_t \tag{6.1}$$

$$\sigma_t^2 = \alpha + \zeta VIX_t + \varepsilon_t \tag{6.12}$$

The table below confirms the above findings. As can be seen in almost all the cases while modelling using only VIX alone relatively high R^2's are reported, however in the majority of the cases it does not outperform the GARCH types forecasts for which higher in value R^2 's are found.

Table 6.13

R^2 /Model	GARCH	TGARCH	EGARCH	APARCH	CGARCH	IGARCH	VIX AR(1)
UK	0.196733	0.250620	0.262685	0.246791	0.199565	0.191066	0.194836
FRANCE	0.158798	0.211450	0.220740	0.207078	0.160721	0.153394	0.186285
USA	0.216410	0.255646	0.271440	0.258287	0.223984	0.197869	0.231813

6.7 Conclusion

After evaluating the performance of GARCH type models by producing one step ahead volatility forecasts for each of the three markets namely the UK, France and the USA, the explanatory power of VIX and Trading Volume is assessed and compared, within a MZ regression framework. The results show that the models perform slightly better when the VIX and Volume are included in any type of specification adopted, however the combination of VIX and Volume produce even better results. We thus provide evidence that VIX and Volume have an explanatory power in tracking the value of a market based volatility.

When assessing the significance of the added components and their explanatory power through a forecast encompassing exercise it is revealed that both VIX and Volume in the majority of cases have positive and significant coefficients suggesting that both VIX and Volume carry additional information and are not encompassed by the GARCH type forecast. However, although statistically their added value is confirmed, the coefficients are very small, and most cases close to zero especially for Volume, suggesting that economically their added value if of little importance.

The final question we attempt to answer is whether VIX alone can be used to more accurately forecast volatility than the GARCH type models. Modelling VIX as a basic AR(1) model further encompassing tests are performed reaching the same as before conclusion for all the three markets that GARCH type forecasts are adequate and although one cannot ignore the added value of including VIX -by increasing the R^2 values, the coefficients are too small that the superiority of the GARCH type models is confirmed.

7. Summary and conclusions

7.1 Summary

In the literature review chapter definitions, grounding theories and future research frontiers set the foundations for the next chapters. This thesis enters the ongoing debate in identifying the most suitable and accurate models and methodologies with the aim of producing accurate volatility forecasts. Even though this is not a new topic in the finance literature, so far no generally accepted conclusion has been reached.

We enter the debate by carrying out a straight forward volatility exercise in Chapter 3. The forecast ability of a number of representative models belonging to two main model groups are compared. From the 'simple' models two representative techniques are chosen; the Exponential Smoothing and the Moving Averages method and from the more 'advanced' GARCH type models capturing the features of volatility clustering (GARCH (1,1)), the leverage effect (TGARCH and EGARCH) and volatility persistence (CGARCH and HYGARCH), attributes found to exist in the data.

The results show that the more 'advanced' GARCH type models perform better overall than the 'simple' models. Ranking the models by performance we see that the asymmetric models come first followed in second place by the long memory models and in third place the 'simpler' time series models are positioned. When taking into account the country classification the ranking process becomes clearer for the developed economies and less clear for the emerging economies when trying to

identify the winner of the exercise. However, when country classification is taken into account for both developed and emerging economies there is no contest in identifying the worst performing model; the Exponential Smoothing model.

The results of this chapter are of interest to anyone involved in the modelling and forecasting of volatility. More specifically, our results confirm the work of the pro GARCH genre supporters suggesting that the more complex models capturing the different stylised effects found in data sets are more accurate in the forecasting process. From a practitioner's point of view the above findings are also important since questions are raised about the suitability of the Exponential Smoothing model, a model widely used by finance practitioners, which ranked low in the exercise and was found to be the worst performing model overall. In addition to the above main findings a further conclusion of importance to academics, practitioners and policymakers, is reached; caution needs to be exercised when attempting to model and forecast the volatility of emerging markets. It becomes less clear which is the most appropriate model to use for the purpose of forecasting volatility, even though it has been confirmed that ARCH effects exist in emerging market data. Other important factors such as financial liberalisation and sensitivity of information flow need to be accounted for.

In continuing the search with the aim of identifying key parameters for improving the accuracy of volatility forecasts Chapter 4 explores a previously unanswered question. 'What is the optimal in-sample length for producing accurate out-of-sample forecasts?' Following on from the previous chapter, the better performing models are put through a backward recursion exercise seeking to explore the two main differing

views on this matter. On one hand small in-sample periods (small number of observations) are chosen by practitioners and investors due to cost and storage restrictions, and on the other hand large in-sample periods (large number of observations) are used by researchers and academics when forecasting volatility. Our findings show a degree of homogeneity where for most markets of our sample and for most models, large in-sample periods are not necessary for producing accurate forecasts, although for the most accurate forecasts larger in-sample periods are used. This result supports the view of the practitioners and investors. Moreover, a pattern emerges when the results are categorised on country classification; smaller in-sample durations are required for producing accurate forecasts in emerging markets but more accurate forecasts produced for countries in developed economies.

The findings of Chapter 4 add a further dimension to the on going debate of how volatility forecasts can be improved and are of significant importance to finance professionals, investors, policymakers and academics. Overall, the view of the practitioners and investors was found to dominate in most cases, but caution needs to be exercised because in several cases there appeared to be a trade off between accuracy and in-sample length. Furthermore, the identified pattern based on market classification is a step forward and an important factor worth considering in future volatility forecasting exercises.

The objective of Chapter 5 was to assess whether the in-sample superiority of the GARCH genre carries over to out-of-sample forecasting, or whether forecasts from the RiskMetrics model, preferred by finance professionals can provide adequate forecasts of volatility in a risk management setting and more specifically in a Value-

at-Risk setting. The RiskMetrics model is known for its simplicity of application, however, the academic finance literature has highlighted the problems associated with the model especially with respect to the undefined unconditional variance and the model's inability to produce long-horizon forecasts. Conversely the GARCH genre of models has found support by the finance literature and does not suffer from the same problems but it is also proven that it is better able to capture the inherent time-dependency within volatility.

After updating the sample used in the previous chapters, thirty-one international stock markets were included and RiskMetrics forecasts were compared to those of the GARCH type models (GRACH(1,1), EGARCH, APARCH, IGARCH, FIGARCH and HYGARCH models) within a 1% and 5% VaR framework. The conclusions of the chapter are summed up as follows: When forecasting at the 1% VaR the RiskMetrics model does a poor job and is in most cases the worst performing model, on the other hand the GARCH type models and more specifically the APARCH model is proven to be superior. On the other hand when forecasting at the 5% VaR then the RiskMetrics model performs adequately. Taking into account the country classification it is found that the RiskMetrics performs well in small emerging markets. These results have implications that are of benefit to both the academic community and finance professionals. First, confirming previous findings on the usefulness of the RiskMetrics model in volatility forecasting exercises and second in demonstrating where the RiskMetrics approach performs well.

The final empirical chapter assesses the effect of two more parameters, namely the Volatility Index (VIX) and Trading Volume, on volatility forecasting. The appealing

properties of VIX and Volume have been recognised by the finance literature, and both these factors have been considered in previous forecasting exercises but mainly individually. Only a very small number of recent studies assessing the impact of both VIX and Volume together on volatility forecasting have been published.

VIX is a forward looking measure defined as a benchmark of expected short-term market volatility upon which futures and options contracts on volatility can be written. On the other hand Trading Volume is caused by information flow which is positively correlated to price changes and volatility. Using a relatively small dataset due to data availability, VIX and Volume data are included within a GARCH type model framework where the testing procedure of Mincer-Zarnowitz (MZ) is followed, and then forecast encompassing tests are also performed in order to establish whether there is added value in incorporating the two parameters within the forecasting process. The findings suggest that both VIX and Volume improve on the informational content of the GARCH type models, VIX does a better job in this process than Volume, but when VIX and Volume are combined, better results are reported. In answering the question whether VIX produces better forecasts than the GARCH genre of models, the answer is no but the informational content of VIX cannot be ignored. As before, the conclusions of this chapter add to the existing literature on volatility forecasting with implications for academics and finance practitioners. VIX and Trading Volume data are easily available and we provide evidence that VIX and Volume have an explanatory power in tracking the value of a market based volatility, on the other hand there is also evidence that GARCH type forecasts are superior.

7.2 Conclusions

This thesis aimed at adding knowledge to the literature on volatility forecasting by means of empirical investigation.[53] The employment of a number of key parameters for improving the accuracy of volatility forecasting were proposed and four main contributions are identified. First, the findings of a volatility forecasting model comparison revealed that the GARCH genre of models are superior compared to the more 'simple' models and models preferred by practitioners. Second, building from the previous findings and with the use of backward recursion forecasts we identified the appropriate in-sample length for producing accurate volatility forecasts, a parameter considered for the first time in the finance literature. Third, further model comparisons were conducted within a Value-at-Risk setting between the RiskMetrics model preferred by practitioners, and the more complex GARCH type models, deriving similar conclusions as before, that GARCH type models are dominant. Finally, two further parameters, the Volatility Index (VIX) and Trading Volume, are considered and their contribution is assessed in the modelling and forecasting process of a selection of GARCH type models. We discover that although accuracy is improved upon, GARCH type forecasts are still superior.

In addition to the four main contributions mentioned above some further aspects are also examined. With the exception of the last empirical chapter where due to data availability only a small number of countries are considered, for the rest of the empirical chapters the samples used are of large datasets from many countries consisting of a good mix of both developed and emerging economies, with the aim of

[53] In terms of research outcome, so far aspects of Chapters 3 and 5 have been published in a refereed journal article. The intention is Chapters 4 and 6 which are currently working papers, is to be presented at conferences and then submitted to journals.

identifying any patterns or other regularities. Furthermore, the nature of the topic allowed for comparisons of methods mainly used in academia and methods preferred by finance practitioners.

The GARCH genre of models are a reduced-form approach and as such have little or no economic theory behind them. Their appealing attribute is that they capture systematic features of the data; especially volatility clustering, asymmetry and long-memory with limited economic meaning to rely on. On the other hand, this class of models has been recognised in the finance community and for which the 2003 Nobel Prize was awarded to their creator Robert F. Engle and is still the most recent Nobel Prize in Economics for a financial market related achievement.

The above mentioned regularities (volatility clustering, the leverage effect and thick tails) arise due to the rate of information flow –news arrival, and how investors react to the information flow (clustering), examples of such theories are the Mixture of Distribution Hypothesis (MDH) or Sequential Information Hypothesis (SIH) and the Noise Trading Hypothesis (NTH). Similarly, information asymmetry can be explained through leverage, feedback and even behavioural explanations exist. The results from the thesis aid in choosing the appropriate model for volatility in recognition of the importance of volatility in pricing and timing decisions.

This thesis enters the ongoing debate in identifying the most suitable and accurate models and methodologies with the aim of producing accurate volatility forecasts. The results show that the more ‘advanced’ GARCH type models perform better overall than the ‘simple’ models. These findings are of interest to anyone involved in

the modelling and forecasting of volatility. From a practitioner's point of view alternative models are suggested, since questions are raised about the suitability of for example, the Exponential Smoothing model, a model widely used by finance practitioners. In addition, caution needs to be exercised when attempting to model and forecast the volatility of emerging markets when the previous picture becomes less clear (it is not so clear identifying the best performing model) for which other important factors such as financial liberalisation and sensitivity of information flow need to be accounted for. A further direct comparison between the Risk Metrics approach preferred by practitioners and the GARCH genre models within a Value at Risk framework lead to a similar conclusion. Risk Metrics does comparatively a poorer job to the GARCH models when a higher degree of accuracy is required. Sample classification again is important since Risk Metrics performs well in small emerging markets.

A question not yet addressed in the finance literature is the optimal in-sample length for producing accurate volatility forecasts. It is found that the view of the practitioners and investors is supported concluding that large in-sample periods are not necessary when forecasting. On the other hand caution needs to be exercised because on several occasions there appeared to be a trade off between accuracy and in-sample length. As before market classification is an important factor worth considering in similar volatility forecasting exercises.

Finally, the appealing properties of VIX and Volume have been recognised by finance practitioners. VIX is used as a forward looking measure setting the benchmark of expected short-term market volatility upon which futures and options contracts on

volatility can be written and Trading Volume is caused by information flow and is positively correlated to price changes and volatility. After performing a further forecasting exercise the superiority of the GARCH models was confirmed however, the value of both VIX and Trading Volume which can easily be obtained cannot be ignored in volatility forecasting exercises.

The importance of volatility forecasting can be realised by the vast literature that exists making it one of the most popular topics in finance. Conversely there still exist many unanswered questions and areas which require further research. As in all empirical investigations there are several methodological aspects that can be criticised which are not always easy to advocate. Model selection, data frequency, sample selection or even the type of software package used are factors not always easy to defend and could also form the basis of further research on the topic. Nevertheless, there are few areas which could form the basis for further research from which the finance literature and the finance practitioner and investment community would benefit from.

Throughout this thesis daily data were used in empirical investigations. Data frequency is an important factor in the volatility forecasting domain with more recent studies using intra daily and high frequency observations. An aspect worth further exploration is the effect of data frequencies on the accuracy of volatility forecasting in an attempt to answer questions of the type: ‘do intra daily observations provide better daily forecasts’ or ‘do daily observations provide better weekly forecasts’? In Chapter 4 of this thesis we addressed the question of identifying suitable in-sample lengths for one step ahead volatility forecasts. Based on the underlying principles and the

methodology followed in this chapter a further question is raised: ‘how far ahead can we forecast’? As can be seen the topic of volatility forecasting is a topic with ongoing research potential.

8. Bibliography and references

Abu Hassan Shaari Mohd, N., and Chin Wen, C., (2007), "An empirical study of realized volatility and trading volume dynamics", *International Research Journal of Finance and Economics*, Vol. 9, pp. 160-166.

Admati, A., and Pfleiderer, P., (1988), "A Theory of Intraday Patterns: Volume and Price Variability", *Review of Financial Studies*, Vol. 1, pp. 3-40.

Aggarwal, R., Inclan, C., and Leal, R., (1999), "Volatility in emerging stock markets", *Journal of Financial and Quantitative Analysis*, Vol. 34, pp.33-55.

Ahoniemi, K., (2008), "Modelling and Forecasting the VIX Index", Unpublished working paper, Helsinki School of Economics and HECER.

Ahoniemi, K., and Lanne, M., (2007), "Joint Modelling of Call and Put Implied Volatility", HECER Discussion Paper, 198.

Akgiray, V., (1989), "Conditional heteroscedasticity in the time series of stock returns: evidence and forecasts", *Journal of Business*, Vol. 62, pp. 55-80.

Alford, A., and Boatsman, J., (1995), "Predicting long-term stock return volatility: Implications for accounting and valuation of equity derivatives", *Accounting Review*, Vol. 70, pp. 599-618.

Alper, C., and Yilmaz, K., (2004), "Volatility and contagion: evidence from the Istanbul stock exchange", *Economic Systems*, Vol. 28, pp. 353-367.

Andersen, T. G., (1996), "Return volatility and trading volume: an information flow interpretation of stochastic volatility", *Journal of Finance*, Vol. 51 (1), pp.169-204.

Andersen, T. G., and Bollerslev, T., (1998), “Answering the skeptics: yes, standard volatility models do provide accurate forecasts”, *International Economic Review*, Vol. 39, pp. 889-905.

Andersen, T. G., and Bollerslev, T., (1997), “Answering the critics: Yes, ARCH models do provide good volatility forecasts”, *NBER Working Paper Series*, Working paper 6023.

Andersen, T. G., Bollerslev, T., Christoffersen, P. F., and Diebold, F. X., (2005), “Volatility Forecasting”, PIER Working Paper (05-011).

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P., (2001), “The distribution of exchange rate volatility”, *Journal of the American Statistical Association*, Vol. 93 (453), pp. 42-55, correction published in 2003, volume 98, page 501.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P., (2000), “Great realisations”, Risk, Vol. 13, pp.105-108.

Andersen, T. G., Bollerslev, T., and Lange, S., (1999), “Forecasting financial market volatility: Sample frequency vis-à-vis forecast horizon”, *Journal of Empirical Finance*, Vol. 6, pp. 457-477.

Andersen, T., Dobrev, D., and Schaumburg, E., (2008), Duration-based volatility estimation. Working paper

Antoniou, A., Ergul, N., Holmes, P., and Priestley, R., (1997), “The impact of regulatory changes on stock market volatility and the cost of equity capital: Evidence

from an emerging market", *European Financial Management*, Vol. 3, (2), pp. 175-190.

Antoniou, A., and Holmes, P., (1995), "Futures trading, information and spot price volatility: evidence for the FTSE-100 Stock Index Futures contrast using GARCH", *Journal of Banking and Finance*, Vol. 19, pp.117-129.

Appiah-Kusi, J., and Menyah, K., (2003), "Return predictability in African stock markets", *Review of Financial Economics*, Vol. 12, pp. 247-270.

Appiah-Kusi, J., and Pescetto, M. G., (1998), "Volatility and Volatility Spill-Overs in Emerging Markets: "The Case of the African Stock Markets", *Ekonomia. The Journal of the Cyprus Economic Society.* Vol. 2, Winter 1998, pp.171-185.

Arestis, P., and Demetriades, P., (1997), "Financial Development and Economic Growth: Assessing the Evidence", *The Economic Journal*, Vol. 107, pp. 783-799.

Balaban, E., Bayar, A., and Faff, R. W., (2006), "Forecasting stock market volatility: further international evidence", *The European Journal of Finance*, Vol. 12, pp. 171-188.

Baillie, R. T., and Bollerslev, T., (1992), "Prediction in Dynamic Models with Time Dependent Conditional Variances", *Journal of Econometrics*, Vol. 52, pp. 91-113.

Baillie, R., T., and Bollerslev, T., (1991), "Intra Day and Inter Day Volatility in Foreign Exchange Rates", *Review of Economic Studies*, Vol. 58, pp. 565-585.

Baillie, R. T., and Bollerslev, T., (1989), "The Message in Daily Exchange Rates: A Conditional Variance Tale", *Journal of Business and Economic Statistics*, Vol. 7, pp.297-305.

Baillie, R. T., Bollerslev, T., and Mikkelsen, H. O., (1996), "Fractionally Integrated Generalised Autoregrssive Conditional Heteroskedasticity", *Journal of Econometrics*, Vol. 74 (1), pp. 3-30.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N., (2008), "Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise", *Econometrica*, Vol. 76, pp. 1481-1536.

Barndorff-Nielsen, O. E., and Shephard, N., (2004), "Power and bipower variation with stochastic volatility jumps", *Journal of Financial Econometrics*, Vol. 2, pp.1-48.

Barndorff-Nielsen, O. E., and Shephard, N., (2002), "Econometric analysis of realised volatility and its use in estimating stochastic volatility models", *Journal of the Royal Statistical Society*, Series B 64, pp. 253-280.

Bauwens, L., and Storti, G., (2007), "A Component GARCH model with Time Varying Weights", Studies in Nonlinear Dynamics & Econometrics, 13/2, article 1, 2009. Revision of CORE DP 2007/19.

Bauwens, L., Preminger, A., and Rombouts, J. V. K., (2006), "Regime switching GARCH models", CORE Discussion Paper, 2006/11.

Beckers, D., (1981), "Standard Daviations Implied in Option Prices as Predictors of Future Stock Price Variability", *Journal of Banking and Finance*, Vol. 5, pp. 363-381.

Becker, R., Clements, A. E., and White, S. I., (2006), "On the Informational Efficieincy of S&P500 Implied Volatility", *North American Journal of Economics and Finance*, Vol. 17, pp. 139-153.

Bekaert, G., and Harvey, C., R. (2002), "Research in emerging markets: looking to the future", *Emerging Markets Review*, Vol. 3, pp. 429-448.

Bekaert, G., and Harvey, C. R., (1997), "Emerging equity market volatility". *Journal of Financial Economics*, Vol. 43, (1), pp.29-77.

Bera, A. K., and Higgins, M. L., (1993), "ARCH models: Properties, estimation and testing", *Journal of Economic Surveys*, Vol. 7 (4), pp. 305-362.

Bessembinder, H., and Seguin, P. J., (1993), "Price Volatility, Trading Volume, and Market Depth: Evidence from Futures Markets", *Journal of Financial and Quantitative Analysis*, Vol. 28 (1), pp. 21-39.

Black, F., (1976), "Studies of Stock Price Volatility Changes", *Proceedings from the American Statistical Association, Business and Economic Statistic Section*, pp. 177-181.

Black, F., and Scholes, M., (1973), "The Pricing of Options and Corporate Liabilities", *Journal of Political Economy*, Vol. 81, pp. 637-654.

Blair, B., Poon, S. H., and Taylor, S. J., (2001), "Forecasting S&P 100 Volatility: The Incremental Information Content of Implied Volatilities and High Frequency Index Returns", *Journal of Econometrics*, Vol. 105, pp. 5-26.

Blattberg, R. C., and Gonedes, N. J., (1974), "A Comparison of the Stable and Student Distribution of Statistical Models for Stock Prices", *Journal of Business*, Vol. 47, pp. 244-280.

Bollerslev, T., (1986), "Generalised Autoregressive Conditional Heteroskedasticity". *Journal of Econometrics*, Vol. 31, pp. 307-327.

Bollerslev, T., Engle, R. F., and Nelson, D. B. (1994), "ARCH models", *Handbook of Econometrics*, Vol. 4, Elsevier Science.

Bollerslev, T., Engle, R. F., and Wooldridge, J. M., (1988), "A capital asset pricing model with time-varying covariance", *Journal of Political Economy*, Vol. 96, (1), pp. 116-131.

Bollerslev, T., Chou, R. Y., and Kroner, K. F., (1992), “ARCH modelling in Finance”, *Journal of Econometrics*, Vol. 52, pp. 5-59.

Bollerslev, T., and Jubinski, D., (1999), ”Equity Trading Volume and Volatility: Latent Information Arrivals and Common Long-Run Dependencies”, *Journal of Business and Economic Statistics*, Vol. 17 (1), pp. 9-21.

Bollerslev, T. and Mikkelsen, H. O., (1996), “Modelling and Pricing Long-Memory in Stock Market Volatility,” *Journal of Econometrics, Vol.* 73, pp. 151-184.

Boudoukh, J., Richardson, M., and Whitelaw, R. F., (1997), “Investigation of a Class of Volatility Estimators”, *Journal of Derivatives*, Vol. 4, pp. 63-71.

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

Brock, W. A., and LeBaron, B. D., (1996), “A dynamic structural model for stock return volatility and trading volume”, *Review of Economics and Statistics*, Vol. 78 (1), pp. 94-110.

Brooks, C., (2008), “Introductory econometrics for finance”, Second edition UK: Cambridge Press.
Brooks, C., (2002), “Introductory econometrics for finance”, First edition, UK: Cambridge Press.

Brooks, C., (1998), “Predicting stock index volatility: Can Market Volume help?”, *Journal of Forecasting* Vol. 17, pp. 59–98.

Brooks, R., (2007), “Power arch modelling of the volatility of emerging equity markets”, *Emerging Markets review*, Vol. 8, pp. 124-133.

Brooks, R., Davidson, S., and Faff, R., (1997), “An evaluation of the effects of major political change on stock volatility: the South African experience”, *Journal of International Financial Markets, Institutions and Money*, Vol. 7, pp. 255-275.

Campbell, S. D., (2005), “A Review of Backtesting and Backtesting Procedures”, Working Paper, Federal Reserve Board.

Campbell, J. Y., Grossman, S. J., and Wang, J., (1993), “Trading volume and serial correlation in stock returns”, *Quarterly Journal of Economics*, Vol. 108, pp. 905-935.

Campbell, J. Y., Lo, A. W., and MacKinlay, A. C., (1997), “The Econometrics of Financial Markets”, USA: Princeton University Press.

Canina, L., and Figlewski, S., (1993), “The informational Content of Implied Volatility”, Review of Financial Studies, Vol. 6, pp. 659-681.

Carr, P., and Wu, L., (2006), “A Tale of Two Indices”, *Journal of Derivatives*, Vol. 13 (3), pp. 13-29.

Chang, C., Jimenez-Martin, J. A., McAleer, M., and Perez-Amaral, T., (2011), “Risk Management of Risk under the Basel Accord: Forecasting Value at Risk of VIX Futures”, Working Paper. http://ssrn.com/abstract=1765202

Chiras, D. P., and Manaster, S., (1978), “The information content of option prices and a test for market efficiency”, *Journal of Financial Economics*, Vol.6, pp. 213-234.

Chong, W. C., Ahmad, M. I., and Abdullah, M. Y., (1999), “Performance of GARCH Models in Forecasting Stock Market Volatility”, *Journal of Forecasting* Vol. 18, pp. 333-343.

Chong, T. Y., and Hendry, D. F., (1986), “Econometric evaluation of linear macroeconomic models”, *Review of Economic Studies*, Vol. 53, pp. 671-690.

Chou, R. Y., (1988), “Volatility Persistence and Stock Valuation: Some Empirical Evidence using GARCH”, *Journal of Applied Econometrics*, Vol.3, pp.279-294.

Choudhry, T., (1996), “Stock Market Volatility and the crash of 1987: Evidence from six emerging markets”, *Journal of International Money and Finance.* Vol. 15, (6), pp. 969-981.

Christensen, B. J., Nielsen, M. O., and Zhu, J., (2010), “Long memory in stock market volatility and the volatility-in-mean effect: The FIEGARCH-M Model”, *Journal of Empirical Finance*, Vol. 17, pp. 460-470.

Christensen, B. J., and Prabhala, N. R., (1998), “The Relation between Implied and Realised Volatility”, *Journal of Financial Economics*, Vol. 50 (2), pp. 125-50.

Christensen, B. J., Hansen, C. S., and Prabhala, N. R., (2001), “The Telescoping Overlap problem in Options Data”, School of Economics and Management, University of Aarhus (unpublished manuscript).

Christensen, B. J., and Hansen, C. S., (2002), “New Evidence on the Implied-Realized Volatility Relation”, *European Journal of Finance*, Vol. 8 (2), pp.187-205.

Christiansen, C., and Hansen, C. S., (2002), “Implied Volatility of Interest Rate Options; An Empirical Investigation of the Market Model”, *Review of Derivatives Research*, Vol. 5 (1), pp. 51-80.

Christie, A. A., (1982), “The Stochastic Behavior of Common Stock Variances: Value, Leverage and Interest Rate Effects”, *Journal of Financial Economics*, Vol. 10, pp. 407-432.

Chu, C. S. J., (1995), “Detecting parameter shift in GARCH models” *Econometric Reviews*, Vol. 14, pp. 241-266.

Chu, S. H., and Freund, S., (1996), “Volatility Estimation For Stock Index Options: A GARCH Approach”, *Quarterly Review of Economics and Finance*, Vol. 36 (4), pp. 431-450.

Clark, P. K., (1973), “A subordinated stochastic process model with finite variance for speculative prices”, *Econometrica*, Vol. 41, pp. 135-155.

Cochrane, J. H., (1991), “Volatility tests and efficient markets. A review essay”, *Journal of Monetary Economics*, Vol. 27, (3), pp. 463-485.

Copeland, T. E., (1976), “A model of asset trading under the assumption of sequential information arrival”, *Journal of Finance*, Vol. 31, pp. 1149-1168.

Cornell, B., (1978), “Using the Options Pricing Model to Measure the Uncertainty Producing Effect of Major Announcements”, *Financial Management*, Vol. 7, pp. 54-59.

Corrado, C. J., and Miller, T. W., (2005), “The Forecast Quality of CBOE Implied Volatility Indexes”, *The Journal of Futures Markets*, Vol. 25, pp. 339-373.

Crouhy, M., Galai, D., and Marl, R., (2006), “The Essentials of Risk Management”, McGraw-Hill, New York, USA.

Cumby, R., Figlewski, S., and Hasbrouck, J., (1993), “Forecasting volatilities and correlations with EGARCH models”, *Journal of Derivatives*, Vol. 1, pp. 51-63.

Cunado, J., Biscarri, J., and Gracia, F., (2006), “Changes in the dynamic behaviour of emerging market volatility: revisiting the effects of financial liberalisation”, *Emerging Markets review*, Vol. 7, pp. 261-278.

Daal, E., Naka, A., and Yu, J. S., (2004) “Volatility clustering, leverage effects, and jumps dynamics in emerging Asian equity markets” *Department of Economics and Finance Working Papers 1991-2006*., Paper 25, http://scholarworks.uno.edu/econ_wp/25

Damodaran, A., “Value at Risk”, http://pages.stern.nyu.edu/~adamodar/

Damodaran, A., "Risk Management: A Corporate Governance Manual", http://pages.stern.nyu.edu/~adamodar/

Daniels, J. D., and Radebaugh, L. H., (1998), "International Business Environments and Operations", Eighth edition. USA: Addison Wesley.

Davidson, J., (2004), "Moment and Memory Properties of Linear Conditional Heteroscedasticity Models, and a New Model", *Journal of Business and Economic Statistics*, Vol. 22 (1), pp. 16-29.

De Long, J. B., Shleifer, A., Summers, L. H. and Waldmann, R. J., (1990), "Positive Feedback Investment Strategies and Destabilizing Rational Speculation", *Journal of Finance*, Vol. 45, (2), pp. 379-395.

De Santis, G., and Imrohoroglu, S., (1997), "Stock returns and volatility in emerging financial markets", *Journal of International Money and Finance*, Vol. 16, pp. 561-579.

Dempster, M. A. H., (2002), "Risk Management: Value at Risk and Beyond", Cambridge University Press: UK.

Dennis, P., Mayhew, S., and Stivers, C., (2006), "Stock Returns, Implied Volatility Innovations, and the Asymmetric Volatility Phenomenon", *Journal of Financial and Quantitative Analysis*, Vol. 41, pp. 381-406.

Diebold, F. X., and Lopez, J., (1996), "Forecast evaluation and combination. In: Maddala, G. S., Rao, C. R. (Eds), Handbook of Statistics. North-Holland, Amsterdam, pp. 241-268.

Diebold, F. X., and Nerlove, M., (1989), "The Dynamics of Exchange Rate Volatility: A Multivariate Latent Factor ARCH Model", *Journal of Applied Econometrics*, Vol. 4, pp. 1-21.

Dimitrakopoulos, D. N., Kavussanos, M. G., and Spyrou, S. I., (2010) “Value at risk models for volatile emerging market equity portfolios”, *The Quarterly Review of Economics and Finance*”, Vol. 10, pp. 515-526.

Dimson, E., and Marsh, P., (1990), “Volatility forecasting without data snooping”, *Journal of Banking and Finance*, Vol. 14, pp. 399-421.

Ding, Z., and Granger, C. W. J., (1996), “Modelling volatility persistence of speculative returns: A new approach”, *Journal of Econometrics*, Vol. 73 (1), pp. 185-215.

Donaldson, G. R., and Kamstra, M. J., (2005), "Volatility forecasts, trading volume and the ARCH versus option-implied volatility trade-off", *The Journal of Financial research*, Vol. 28 (4), pp. 519-538.

Ding, Z., Granger, C. W. J., and Engle, R. F., (1993), “A long memory property of stock market returns and a new model”, *Journal of Empirical Finance*, Vol. 1, pp. 83-106.

Doyle, R., Scott, D., and Crimmins, C., (1999), “Was The Tequila Effect Rational?”, http://www.economics.tcd.ie/SER/1999/essay14.html

Duffie, D., and Pan, J., (1997), “An Overview of Value at Risk”, *The Journal of Derivatives*, Vol. 4 (3), pp. 7-49.

Ederington, L. H., and Guan, W., (2002), “Is Implied Volatility an Informational Efficient and Effective Predictor of Future Volatility?”, Journal of Risk, Vol. 4, pp. 29-46.

Engle, R. F., (2003), “Risk and volatility: Econometric models and financial practice”. Nobel Lecture, December 8 2003, New York University, USA.

Engle, R. F., (2002), “New frontiers for ARCH models”, *Journal of Applied Econometrics*”, Vol. 17, pp.425-446.

Engle, R. F., (1982), "Autoregressive Conditional Heteroscedasticity with estimates of the variance of UK inflation", *Econometrica*, Vol. 50, pp. 987-1008.

Engle, R. F., and Bollerslev, T., (1986), "Modelling the persistence of conditional variances", *Econometric Reviews*, Vol. 5, pp.1-50.

Engle, R. F., and Gallo, G. M., (2006), A multiple indicators model for volatility using intra-daily data, *Journal of Econometrics*, Vol. 131, pp.3-27.

Engle, R. F., and Kroner. F. K., (1995), "Multivariate Simultaneous Generalized ARCH", *Econometric Theory*, Vol. 11, pp.122-150.

Engle, R. F., and Lee, G. G. J., (1999), "A permanent and transitory component model of stock return volatility", *In Engle, R F and White H (eds) Cointegration, Causality and Forecasting: A Festschrift in Honour of Clive W J Granger*, Oxford, UK: Oxford University Press.

Engle, R. F., Lilien, D., and Robins, R., (1987) "Estimating time varying risk premia in the term structure: The ARCH-M model", *Econometrica*, Vol. 55, pp. 391-407.

Engle, R. F., and Manganelli, S., (2004), CAViaR: Conditional autoregressive value at risk by regression quantiles", *Journal of Business and Economic Statistics*, Vol. 22, pp. 367-381.

Engle, R. F., and Ng, V., (1993), "Measuring and testing the impact of news on volatility", *Journal of Finance*, Vol. 48, pp. 1749-1778.

Engle, R. F., Ng, V., and Rothschild, M., (1990), "Asset pricing with a Factor ARCH Covariance Structure: Empirical Estimates for Treasury Bills", *Journal of Econometrics*, Vol. 45, pp. 213-238.

Engle, R. F., and Patton, A. J., (2001), "What Good is a Volatility Model?" Archive at NYU: http://archive.nyu.edu/handle/2451/26881

Engle, R. F., and Susmel, R., (1993), "Common Volatility in International Equity Markets", *Journal of Business and Economic Statistics*, Vol. 11, pp. 167-176.

Epps, T. W., (1975), "Security price changes and transaction volumes: theory and evidence", *American Economic Review*, Vol. 65, pp. 586-597.

Epps, T. W., and Epps. M. L., (1975), "The stochastic dependence of security price changes and transaction volumes: implications for the mixture of distribution hypothesis", *Econometrica*, Vol. 44, pp. 305-321.

Ernst, C., Stange, S., and Kaserer, C., (2009), 'Measuring Market Liquidity Risk – Which Model Works Best?', Working Paper, Centre for Entrepreneurial and Financial Studies.

EViews User's Guide, USA: Quantitative Micro Software 1997.

EViews User's Guide, USA: Quantitative Micro Software 2010.

Fama, E. F., (1970), "Efficient capital markets: A review of Theory and Empirical Work", *Journal of Finance*, Vol. 25, pp. 383-417.

Fama, E. F., (1965), "The behaviour of stock market prices", *Journal of Business*, Vol. 38, pp.34-105.

Fama, E. F., (1963), "Mandelbrot and the Stable Paretian Distribution", *Journal of Business*, Vol. 36, pp. 420-429.

Fifield, S. G. M., Lonie, A. A., and Power, D. M., (1998), "A review of Research into Emerging Stock Markets". *Economic Issues*, Vol. 3, (1), March 1998.

Figlewski, S., (1997), "Forecasting Volatility", *Financial Markets, Institutions and Instruments*, Vol. 6, Blackwell Publishers. (Note: final draft 2004)

Fleming, J., (1998), "The quality of market volatility forecasts implied by S&P 100 index option prices", *Journal of Empirical Fin*ance, Vol. 5 (4), pp. 317-345.

Fleming, J., Ostdiek, B., and Whaley, R. E., (1995), "Predicting Stock Market Volatility: A New measure", *Journal of Futures Markets*, Vol. 15 (3), pp. 265-302.

Formari, F., and Mele, A., (2006), "Approximating volatility diffusions with CEV-ARCH models", *Journal of Economic Dynamics & Control*, Vol. 30, pp. 931-966.

Franses, P. H., and Dijk, D., (1996), "Forecasting stock market volatility using (nonlinear) GARCH models", *Journal of Forecasting*, Vol. 15, pp. 229-235.

French, K. R., and Roll, R., (1986), "Stock Return Variances: The Arrival of Information and the Reaction of Traders", *Journal of Financial Economics*, Vol. 17, pp. 5-26.

Frennberg, P., and Hansson, B., (1996), "An Evaluation of Alternative Models for Predicting Stock Volatility: Evidence from a Small Stock Market", *Journal of International Financial Markets Institutions and Money*, Vol. 5, pp.117-134.

Frijns, B., Lehnert, T., and Zwinkels, R., (2011), "A volatility targeting GARCH model with time-varying coefficients", *Journal of Empirical Finance*, Vol.18, pp.522-532.

Fuertes, A., Izzeldin, M., and Kalotychou, E., (2009), "On forecasting daily stock volatility: The role of intraday information and market conditions", *International Journal of Forecasting*, Vo. 25, pp. 259-281.

G@rch 6 Help http://www.core.ucl.ac.be/~laurent/G@RCH/site/default.htm

Gerity, M. S., and Mulherin, J. H., (1992), "Trading Halts and Market Activity: An Analysis of Volume at the Open and Close", *Journal of Finance*, Vol. 47, pp. 1765-1784.

Ghysels, E., Harvey, A., and Renault, E., (1996), "Stochastic Volatility", in Handbook of Statistics: statistical Mehtods in Finance, Vol. 14, Maddala G. R., and Rao, C. R., eds. Amsterdam: Elsevier Science, Ch 5, pp 119-191.

Gil-Diaz, F., "The origin of Mexico's 1994 financial crisis", *The Cato Journal*, Winter 1998, Vol. 17 (3), pp. 303-313, Cato Institute, http://www.cato.org

Giot, P., and Laurent, S., (2007), "The information Content of Implied Volatility in Light of the Jump/Continuous Decomposition of Realised Volatility", *Journal of Futures markets*, Vol. 27, pp. 337-359.

Giraitis, L., Kokoszka, P., and Leipus, P., (2000), "Stationary ARCH Models: Dependence Structure and Central Limit Theorem", *Journal of Time Series Analysis*, Vol. 4, pp. 221-238.

Glosten, L., Jagannathan, R. and Runkle, D., (1993), "On the relation between the expected value and the volatility of nominal excess return on stocks", *Journal of Finance*, Vol. 46, pp. 1779-1801.

Gokcan, S. F., (2000), "Forecasting volatility of emerging stock markets: linear versus non-linear GARCH models", *Journal of Forecasting*, Vol. 19, pp. 499-504.

Grabel, I., (1995), "Assessing the Impact of Financial Liberalisation on Stock Market Volatility in Selected Developing Countries", *Journal of Development Studies*, Vol. 31, pp. 903-917.

Gray, S., (1996), "Modeling the conditional distribution of interest rates as a regime-switching process", *Journal of Financial Economics*, Vol. 42, pp.27-62.

Gujarati, D. N., (1995), "*Basic Econometrics"* Third edition, McGraw-Hill: Singapore

Hakim, A., McAleer, M., and Chan, F., (2009), "Forecasting Portfolio Value at Risk for International Stocks, Bonds and Foreign Exchange", Working paper: http://mssanz.org.au/MODSIM07/papers/31_s8/ForecastingPortfolio_s8_Hakim_.pdf

Hamao, Y., Masulis, R., W., and Ng, V. K., (1990), "Correlation in Price Changes and Volatility Across International Stock Markets", *Review of Financial Studies*, Vol. 3, pp. 281-307.

Hamilton, J. D., and Susmel, R., (1994), "Autoregressive Conditional Heteroscedasticity and Changes in Regime", *Journal of Financial Econometrics*, Vol. 2, pp. 493-530.

Hansen, P. R., (2005), "A test for superior predictive ability", *Journal of Business and Economic Statistics*, Vol. 23, pp. 365-380.

Hansen, P. R., and Horel, G., (2009), "Quadratic variation by Markov chains", Working paper. Available at: http://stanford.edu/people/peter.hansen

Hansen, P. R., Huang, Z., and Shek, H. H., (2010), "Realised GARCH: A Joint Model for Returns and Realised Measures of Volatility", *Journal of Applied Econometrics. Forthcoming*. Available at: http://ssrn.com/abstract=1533475

Hansen, P. R., Kim, J., and Lunde, A., (2003), "Testing for superior predictive accuracy using Ox: A manual for SPA for Ox. Available from Peter Hansen, Stanford University.

Hansen, P. R., and Lunde, A., (2005), "A forecast comparison of volatility models: Does anything beat a GARCH(1,1)?", *Journal of Applied Econometrics*, Vol. 20, pp. 873-889.

Haque, M., Hassan, M., Maroney, N., and Sackley, W., (2004), "An empirical examination of stability, predictability and volatility of Middle Eastern and African stock markets", *Review of Middle East Economics and Finance*, Vol. 2, pp. 19-42.

Harris, L., (1986), “A Transaction Data Study of Weekly and Intradaily Patterns in Stock Returns”, *Journal of Financial Economics*, Vol. 16, pp. 99-117.

Harris, L., (1987), “Transactions data tests of the mixture of distributions hypothesis”, *Journal of Financial and Qualitative Analysis*, Vol. 22, pp. 127-141.

Harris, L., (1986), “Cross-security tests of the mixture of distributions hypothesis”, *Journal of Financial and Qualitative Analysis*, Vol. 21, pp. 39-46.

Harvey, A. C., Ruiz, E., and Shephard, N., (1994), “Multivariate Stochastic Variance Models”, *Review of Economic Studies*, Vol. 61, pp. 247-264.

Harvey, A. C., Ruiz, E., and Sentana, E., (1992), “Unobserved Component Time Series Models with ARCH Disturbances”, *Journal of Econometrics*, Vol. 52, pp. 129-158.

Harvey, C. R., and Huang, R. D., (1992), “Information Trading and Fixed Income Volatility”, *Unpublished manuscript*, Department of Finance, Duke University.

Harvey, C. R., and Huang, R. D., (1991), “Volatility in the Foreign Currency Futures Market”, *Review of Financial Studies*, Vol. 4, pp. 543-569.

Hassan, M., Maroney, N., El-Sady, M., and Telfah, A., (2003), “Country risk and stock market volatility, predictability and diversification in the Middle East and Africa”, *Economic Systems*, Vol. 27, pp. 63-82.

He, H., and Wang , J., (1995), “Differential Information and the Dynamic Bahaviour of Stock Trading Volume”, *Review of Financial Studies*, Vol. 8, pp. 919-972.

Heynen, R. C., and Kat, H. M., (1994), “Volatility Prediction: A Comparison of the Stochastic Volatility GARCH(1,1) and EGARCH(1,1) Models”, *Journal of Derivatives*, Vol. 2, pp. 50-65.

Hiemstra, C. and Jones, J. D., (1994), "Testing for linear and nonlinear Granger causality in the stock price-volume relation", *Journal of Finance*, Vol. 49 (5), pp. 1639-1664.

Hiemstra, C., and Jones, J. D., (1993), "Some information relating to the finite sample properties of the modified Baek and Brock nonlinear Granger causality test", Working Paper, University of Strathclyde and Securities Exchange Commission.

Higgins, M. L., and Bera, A. K., (1992), "A class of ARCH models", *International Economic Review*, Vol. 33 (1), pp. 137-158.

Hol, E., and Koopman, S. J., (2001), "Forecasting the Variability of Stock Index Returns with Stochastic Volatility Models and Implied Volatility", http://www.tinbergen.nl.

Holton, G. A., (2002), "History of Value-at-Risk 1922-1998", Working paper.

Hull, J. and White, A., (1987), "The Pricing of options on Assets with Stochastic Volatilities", *Journal of Finance*, Vol. 42, pp. 281-300.

Hwang, S., and Salmon, M., (2004), "Market stress and herding", *Journal of Empirical Finance*, Vol. 11, pp. 585-616.

Iori, G., (2002), "A microsimulation of traders activity in the stock market: the role of heterogeneity, agents' interactions and trade frictions", *Journal of Economic Behavior and Organisation*, Vol. 49 (2), pp.269-285.

Jayasuriya, S., (2005), "Stock market liberalisation and volatility in the presence of favourable market characteristics and institutions", *Emerging Markets Review*, Vol. 6, pp.170-191.

Jenning, R., Starks, L. and Fellingham, J. C., (1981), "An equilibrium model of asset trading with sequential information arrival", *Journal of Finance*, Vol. 36, pp.143-161.

Jiang G., and Tian, Y., (2005), “Model-Free Implied Volatility and Its Information Content”, Review of Financial Studies, Vol. 18, pp. 1305-1342.

Jorion, P., (2007), “Value at Risk: The New Benchmark for Managing Financial Risk”, 3rd edition, McGraw-Hill, New York, USA.

Jorion, P., (2000), “Value at Risk: The New Benchmark for Managing Financial Risk”, McGraw-Hill, New York, USA.

Jorion, P., (1996), "Risk and turnover in the foreign exchange market", in Franke, J.A., Galli, G., Giovannini, A. (Eds), *The Microstructure of Foreign Exchange Markets*, Chicago University Press, Chicago.

Jorion, P., (1996), “Risk 2: “Measuring the Risk in Value at Risk”, *Financial Analysis Journal*, Vol. 52, (6), pp. 47-56.

Jorion, P., (1995), “Predicting volatility in the foreign exchange market”, *Journal of Finance*, Vol. 50, pp. 381-400.

Kanas, A., and Yannopoulos, A., (2001), “Comparing linear and nonlinear forecasts for stock return”, *International Review of Economics and Finance*, Vol. 10, pp. 383-398.

Karpoff, J. M., (1987), “The Relation between Changes and Trading Volume: A survey”, *Journal of Financial and Quantitative Analysis*, Vol. 22, pp. 109-126.

Karpoff, J. M., (1986), “A theory of trading volume”, *Journal of Finance*, Vol. 41, pp. 1069-1082.

Kim, E., and Singal, V., (2000), “The fear of globalising capital markets”, *Emerging Markets review*, Vol. 1, 183-198.

Kim, E. H., and Singal, V., (1993), “Opening Up to Stock Markets by Emerging Economies: Effect on Portfolio Flows and Volatility Stock Prices” in Claessens, S.,

and Gooptu, S., Portfolio Investment in Developing Countries, World Bank Discussion Paper 228 September pp.383-403.

King, M., Sentana, E., and Wadhwani, S., (1994), "Volatility and Links Between National Stock Markets", *Econometrica*, Vol. 78, pp. 901-934.

Klaassen, F., (2002), "Improving GARCH Volatility Forecasts with Regime-Switching GARCH", Empirical Economics, Vol. 27, pp. 363-394.

Kolb, R. W., Becketti, S., and Sellon, G. H., (1991), "Has Financial Market Volatility Increased?", *The investments reader*, Kolb Publishing Company: USA.

Konstantinidi, E., and Skiadopoulos, G., (2011), "Are VIX futures prices predictable? An empirical investigation", *International Journal of Forecasting*, Vol. 27, pp. 543-560.

Konstantinidi, E., Skiadopoulos, G., and Tzagkaraki, E., (2008), "Can the evolution of implied volatility be forecasted? Evidence from European and US implied volatility indices", *Journal of Banking and Finance*, Vol. 32, pp. 2401-2411.

Koutmos, G., and Booth, G. G., (1995), "Asymmetric Volatility Transmission in International Stock Markets", *Journal of International Money and Finance*, Vol. 14, (6), pp.747-762.

Kuester, K., Mittnik, S., and Paolella, M., (2006), "Value at Risk prediction: A comparison of alternative strategies", *Journal of Financial Econometrics*, Vol. 4, pp.53-89.

Kunitomo, N., (1992), "Improving the Parkinson method of estimating security price volatilities", *Journal of Business*, Vol. 65, pp. 295-302.

Kupiec, P., (1995), 'Techniques for verifying the accuracy of risk management models', *Journal of Derivatives*, 2, 73-84.

Lamoureux, C. G., and Lastrapes, W. D., (1994), "Endogenous Trading Volume and Momentum in Stock Return Volatility", *Journal of Business and Economic Statistics*, Vol. 12, pp. 253-260.

Lamoureux, C. G., and Lastrapes, W. D., (1993), "Forecasting Stock-Return Variance: Toward and Understanding of Stochastic Implied Volatility" *Review of Financial Studies*, Vol. 6, pp. 293-326.

Lamoureux, C. G., and Lastrapes, W. D., (1990), "Heteroscedasticity in stock return data: Volume versus GARCH effects", *Journal of Finance*, Vol. 45 (1), pp. 221-229.

Latane. H. A., and Rendleman, R. J., (1976), "Standard deviations of stock price ratios implied in option prices", *Journal of Finance*, Vol. 31, pp.369-381.

Le, V., and Zurbruegg, R., (2010), "The role of trading volume in volatility forecasting", *Journal of International Financial Markets, Institutions & Money*, Vol. 20, pp. 533-555.

Leavens, Dickson H., (1945), 'Diversification of investments', *Trusts and Estates*. 80 (5), pp. 469-473.

LeBaron, B., (1992), "Forecast Improvements using a Volatility Index", *Journal of Applied Econometrics*, Vol. 7, Supplement: Special Issue on Nonlinear Dynamics and Econometrics (Dec., 1992), pp. S137-S149.

Lin, W. L., Engle, R. F., and Ito, T., (1994), Do Bulls and Bears Move Across Borders? International Transmission of Stock Returns and Volatility as the World Turns", *Review of Financial Studies*, Vol. 7, pp. 507-537.

Linter, J., (1965), 'The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets', *Review of Economics and Statistics*, 47, pp. 13-37.

Lofthouse, S., (1996), “Readings in Investments”, First edition, John Wiley & Sons: UK.

Luu, J. C., and Martens, M., (2003), “Testing the Mixture-of-Distributions hypothesis using “Realised” Volatility”, *The Journal of Futures Markets*, Vol. 23 (7), pp. 661-679.

Luukkonen, R., Saikkonen, P., and Terasvirta, T., (1988), “Testing linearity against smooth transition autoregressive models”, Biometrika, Vol. 75, pp. 491-499.

Mandelbrot, B., (1963), “The variation of certain speculative prices”, *Journal of Business*, Vol. 36, pp.394-419.

Manganelli, S., and Engle. R. F., (2001), “Value at Risk models in Finance”, Working Paper No. 75, European Central Bank, Working Paper Series.

Maris, K., Pantou, G., Nikolopoulos, K., Pagourtzi, E., and Assimakopoulos, V., (2004), “A study of financial volatility forecasting techniques in the FTSE/ASE 20 index”, *Applied Economic Letters*, Vol. 11, pp. 453-457.

Markowitz, H. M., (1991), “Portfolio Selection”, First Edition, Blackwell Publishers: US.

Markowitz, H. M., (1952), ‘Portfolio Selection’, *Journal of Finance*, Vol. 7 (1), pp. 77-91.

Melino, A., and Turnbull, M., (1990), “Pricing Foreign Currency Options with Stochastic Volatility”, *Journal of Econometrics*, Vol. 45, pp. 239-265.

McAleer, M., and Caporin, M., (2011), “Ranking multivariate GARCH models by problem dimension: An empirical evaluation”, Kier Discussion Paper Series, Kyoto Institute of Economic Research, Discussion Paper No. 778.

Merton, R. C., (1980), "On Estimating the Expected Return on the Market: An Explanatory Investigation", *Journal of Financial Economics*, Vol. 8, pp. 323-361.

Merton, R. C., (1973), "Theory of Rational Options Pricing", *Bell Journal of Economics and Management Science*, Vo.: 4, pp. 141-183.

McMillan, D. G., and Kambouroudis, D., (2009), "Are RiskMetrics forecasts good enough? Evidence from 31 stock markets", *International Review of Financial Analysis*, Vol. 18, pp. 117-124.

McMillan, D. G., and Speight, A. E. H., (2011), "Daily FX Volatility Forecasts: Can the GARCH(1,1) Model be Beaten using High-Frequency Data?", *Journal of Forecasting*, Forthcoming.

McMillan, D. G., and Speight, A. E. H., (2004), "Daily Volatility Forecasts: Reassessing the Performance of GARCH Models", *Journal of Forecasting*, Vol. 23, pp. 449-460.

McMillan, D. G., and Speight, A., (2002), "Return-volume dynamics in UK futures", *Applied Financial Economics*, Vol. 12 (10), pp. 707-713.

McMillan, D. G., Speight, A., and Ap Gwilym, O., (2000) "Forecasting UK stock market volatility", *Applied Financial Economics*, Vol. 10, pp.435-448.

Mikosch, T., and Starica, C., (2004), "Nonstationarities in Financial Time Series, the Long-Range Dependence, and the IGARCH effects", *Review of Economics and Statistics*, Vol. 86, pp. 378-390.

Mills, T. C., (1994), "The econometric modelling of financial time series", First edition, Cambridge University Press: UK.

Milton, H., and Raviv, A., (1993), "Differences of opinion make a horse race", *The Review of Financial Studies*, Vol. 6 (3), pp. 473-506.

Mincer, J., and Zarnowitz, V., (1969), “The evaluation of economic forecasts”. in J. Mincer (ed), Economic forecasts and expectations, pp 3-46. New York: Columbia University Press.

Mossin, J., (1966), “Equilibrium in a capital asset market”, *Econometrica*, 34, pp. 768-783.

Nanto, D. K., (1998), CRS Report: “The 1997-98 Asian Financial Crisis”, http://www.fas.org/man/crs/crs-asia2.htm.

Nelson, D. B., (1990a), “Stationarity and persistence in the GARCH (1,1) model”, *Econometric theory*, Vol. 6, pp.318-334.

Nelson, D. B., (1990b), “ARCH models and diffusion approximations”, *Journal of econometrics*, Vol. 45, pp. 7-38.

Nelson, D. B., (1991), “Conditional heteroscedasticity in asset returns: a new approach”, *Econometrica*, Vol. 59 (2), pp. 347-370.

Pagan, A. R. and Schwert, G. W. (1990), “Alternative models for conditional stock volatility”, *Journal of Econometrics*, Vol. 45, pp. 267-290.

Patell, J. M., and Wolfson, M. A., (1981), “The Ex-Ante and Ex-Post Price Effects of Quarterly Earnings Announcement Reflected in Option and Stock Price”, *Journal of Accounting Research*, Vol. 19, pp. 434-458.

Patell, J. M., and Wolfson, M. A., (1979), “Anticipated Information Releases Reflected in Call Option Prices”, *Journal of Accounting and Economics*, Vol. 1, pp. 117-140.

Pong, E., Shackleton, M. B., Taylor, S. J., and Xu, X., (2004), “Forecasting currency volatility: a comparison of implied volatilities and AR(FI)MA models”, *Journal of Banking and Finance”*, Vol. 28 (10), pp. 2541-2563.

Poon, S., and Granger, C. W. J., (2005), "Practical issues in Forecasting Volatility", *Financial Analysis Journal*, Vol. 61 (1), pp. 45-56.

Poon, S., and Granger, C. W. J., (2003), "Forecasting Volatility in Financial Markets: A Review", *Journal of Economic Literature*, Vol. 41 (2), pp. 478-539.

Rabemananjara, R., and Zakoian, J. M., (1993), "Threshold ARCH models and asymmetries in volatility", *Journal of Applied Econometrics*, Vol. 8 (1), pp. 31-49.

Richards, A. J., (1996) "Volatility and Predictability in National Stock Markets: How do Emerging and Mature Markets Differ?", IMF Staff Papers, 43 (3). Pp. 461-501.

Richardson, M., and Smith, T., (1994), "A Direct test of the Mixture of Distributions Hypothesis: Measuring the Daily Flow of Information", *Journal of Financial and Quantitative Analysis*, Vol. 29, pp. 101-116.

RiskMetrics (1996), RiskMetrics™ technical document (fourth edition), J.P. Morgan/Reuters (1996).

Ross, S. A., (1989), "Information and Volatility: The No-Arbitrage Martingale Approach to Timing and Resolution Irrelevancy". *Journal of Finance*, Vol. 44. pp.1-17.

Roy, A. D., (1952), 'Safety first and the holding of assets', *Econometrica*, 20 (3), pp.431-449.

Schwert, W. G., (2011), "Stock Volatility during the recent financial crisis", Working Paper (16976).

Schwert, W. G., (1998), "Stock Market Volatility: Ten Years After The Crash", Working Paper (6381).

Schwert, W. G., (1990) "Stock Volatility and the Crash of '87", *The Review of Financial Studies*, Vol. 3, (1), pp.77-102.

Schwert, G. W., (1990), “Stock Market Volatility”, *Financial Analysis Journal*, Vol. 46, pp. 23-34.

Schwert, G. W., (1989), “Why Does Stock Market Volatility Change Over Time?”, *Journal of Finance*, Vol. 44, pp. 1115-1153.

Schwert, W., G., (1989), “Business Cycles, Financial Crises and Stock Volatility”, *Carnegie-Rochester Conference Series on Public Policy*, Vol. 39, pp. 83-126.

Schwert, G. W., and Seguin, P. J., (1990), “Heteroscedasticity in Stock Returns”, *Journal of Finance*, Vol. 45, pp. 1129-1155.

Sentana, E. (1995), “Quadratic ARCH Models,” *Review of Economic Studies*, Vol. 62 (4), 639-661.

Sharpe, W. F., (1964), “Capital asset prices: A theory of market equilibrium under conditions of risk”, *Journal of Finance*, 19 (3), pp. 425-442.

Sharpe, W. F., (1963), “A simplified model for portfolio analysis”, *Management Science*, Vol. 9, pp. 277-293.

Simons, D. P., (2003), “The NASDAQ Volatility Index During and After the Bubble”, *Journal of Derivatives*, Vol. 11 (2), pp. 9-22.

Smirlock, M., and Starks, L., (1984), “A transactions approach to testing information arrival models”, Typescript, University of Washington, St. Louis.

Smith, R. C., and Walter, I., (1996), “Rethinking Emerging Markets”. *Washington Quarterly*, Vol. 19, (1), pp. 45-64.

So, M., and Yu, P., (2006), “Empirical analysis of GARCH models in value at risk estimation”, *Journal of International Financial Markets, Institutions and Money*, Vol. 16, pp. 180-197.

Spyrou, I. S., and Kassimatis, K., (1999), "Did Equity Market Volatility Increase Following the Opening of Emerging Markets to Foreign Investors?", *Journal of Economic Development*, Vol. 24, (1), pp.39-51.

Szakmary, A., Ors, E., Kim, J. K., and Davidson, W., (2003), "The Predictive Power of Implied Volatility: Evidence from 35 Futures Markets", *Journal of Banking and Finance*, Vol. 27, pp. 2151-2175.

Tauchen, G. E., and Pitts, M., (1983), "The price variability-volume relationship on speculative markets", *Econometrica*, Vol. 51, pp. 485-505.

Taylor, N., (2008), "Can idiosyncratic volatility help forecast market volatility?", *International Journal of Forecasting*, Vol. 24, pp. 462-479.

Taylor, S. J., (2004), "Volatility forecasting with smooth transition exponential smoothing", *International Journal of Forecasting*, Vol. 20, pp. 273-286.

Taylor, S. J., (1987), "Forecasting of the volatility of currency exchange rates", *International Journal of Forecasting*, Vol. 3, pp. 159-170.

Taylor, S. J., (1986), "Modelling Financial Time Series", First edition, Wiley & Sons: UK.

Terasvirta, T., (2006), "An Introduction to Univariate GARCH Models", *SSE/EFI Working Papers in Economics and Finance*, No. 646. December 7, 2006.

Timmermann, A., and Granger, C. W. J., (2004), "Efficient market hypothesis and forecasting", *International Journal of Forecasting*, Vol. 20, pp. 15-27.

Tobin, J., (1958), "Liquidity preference as behaviour towards risk", *The Review of Economic Studies*, 25, pp. 65-86.

Treynor, J., (1961), 'Towards a theory of market value of risky assets', Unpublished manuscript.

Tse, Y. K., (1991), "Stock return volatility in the Tokyo stock exchange", *Japan and the World Economy*, Vol. 3, pp. 285-298.

Tse, Y. K., and Tung, S. H., (1992), "Forecasting volatility in the Singapore stock market", *Asia Pacific Journal of Management*, Vol. 9, pp. 1-13.

Valdez, S., (2000), "An introduction to global financial markets", Third edition, Macmillan Business: UK.

Velasco, A., (1993), "Comments on Opening Up of Stock Markets by Emerging Economies: Effects on Portfolio and Volatility Stock Prices" by Kim, E., H., and Singal, V., in Claessens, S., and Gooptu, S., Portfolio Investment in Developing Countries, World Bank Discussion Paper 228 September pp.390-392.

Wagner, N., and Marsh, T. A., (2005), Surprise volume and heteroscedasticity in equity market returns", *Quantitative Finance*, Vol. 5 (2), pp. 153-168.

Walsh, D., M., and Yu-Gen Tsou, G., (1998), "Forecasting index volatility: sampling interval and non-trading effects", *Applied Financial Economics*, Vol. 8, (5), pp. 477-485.

Wang, J., (1994), "A model of competitive stock trading volume", *Journal of Political Economy*, Vol. 10, pp. 127-168.

Whaley, R. E., (2009), "Understanding VIX", *Journal of Portfolio Management*, Vol. 35, pp. 98-105.

Whaley, R. E., (1993), "Derivatives on market volatility: hedging tools long overdue", *Journal of Derivatives*, Vol. 1, pp. 71–84.

White, H., (2000), “A reality check for data snooping”, *Econometrica*, Vol. 68, pp. 1097-1126.

World Bank Organisation (1997) “Emerging Stock Markets Factbook 1997: Key Trends in Emerging Stock Markets in 1996” www.worldbank.org/html/extpb/emfb97ky.htm

9. Appendix

Appendix 1

Table presenting ARCH Effects test. Ho: No ARCH effects					
Country/ Model	**GARCH**	**EGARCH**	**TGARCH**	**CGARCH**	**HYGARCH**
	Reject Ho: Y /N	Reject Ho: Y /N	Reject Ho: Y /N	Reject Ho: Y /N	Reject Ho: Y /N
Australia	N	**Y**	N	N	N
Austria	N	N	N	N	N
Belgium	N	N	N	N	N
Brazil	N	N	N	N	N
Chile	N	**Y**	N	N	N
Denmark	N	N	N	N	N
France	N	N	N	N	N
Germany	N	N	N	N	N
Hong Kong	N	N	N	N	N
India	N	N	N	N	N
Indonesia	N	**Y**	N	N	N
Ireland	N	**Y**	N	N	N
Israel	N	N	N	N	N
Japan	N	N	N	N	N
Korea	N	N	N	N	N
Malaysia	N	**Y**	N	N	N
Netherlands	N	N	N	N	N
Philippines	N	N	N	N	N
Singapore	N	N	N	N	N
Spain	N	N	N	N	N
Sweden	N	N	N	N	N
Thailand	N	N	N	N	N
Turkey	N	N	N	N	N
UK	N	N	N	N	N
USA	N	N	N	N	N

Appendix 2

Table presenting the Autocorrelations of Squared Standardised residuals. Ho: No Autocorrelation					
Country/ Model	**GARCH**	**EGARCH**	**TGARCH**	**CGARCH**	**HYGARCH**
	Reject Ho: Y /N	Reject Ho: Y /N	Reject Ho: Y /N	Reject Ho: Y /N	Reject Ho: Y /N
Australia	N	N	N	N	N
Austria	N	N	N	N	N
Belgium	N	N	N	N	N
Brazil	N	**Y**	N	N	N
Chile	N(not at 5% and 10%)	**Y**	N(not at 5% and 10%)	N(not at 5%)	N
Denmark	N	N	N	N	N
France	N	N	N	N	N
Germany	N	N	N	N	N
Hong Kong	N	N	N	N	N
India	N	N	N	N(not ay 10%)	N
Indonesia	N	N	N	N(not at 5% and 10%)	N
Ireland	N	N	N	N	N
Israel	N	N	N	N	N
Japan	N	N	N	N	N
Korea	N	N	N	N	N
Malaysia	N	N	N	N(not at 10%)	N
Netherlands	N	N	N	N	N
Philippines	N(not at 5% and 10%)	N(not at 5% and 10%)	N(not at 5% and 10%)	N(not at 5% and 10%)	N(not at 5% and 10%)
Singapore	N(not at 10%)	N(not at 10%)	N(not at 10%)	N(not at 10%)	N(not at 10%)
Spain	N	N	N	N	N
Sweden	N	N	N	N	N
Thailand	N(not at 10%)	N(not at 5% and 10%)	N(not at 10%)	N(not at 10%)	N(not at 10%)
Turkey	N(not at 10%)	N	N(not at 10%)	N	N
UK	N	N	N	N	N
USA	N	N	N	N	N

Appendix 3

Table for the MAPE statistic for volatility forecast models for all sample

Model/Country	GARCH	EGARCH	TGARCH	CGARCH	HYGARCH	Exponential Smoothing	Moving Average
Australia	1.165236	1.11588	1.160944	1.141631	0.929185†	1.251073•	1.090129
Austria	2.541562	2.443325	2.478589	2.544081	2.387909†	2.715365•	2.443325
Belgium	1.576835	1.522936†	1.549312	1.577982	1.616972	1.948394•	1.662844
Brazil	5.066477	5.353977•	5.085227	4.896023	1.886364†	5.198295	2.535795
Chile	1.173052	1.186262	1.178336	1.137384†	1.190225	1.46103•	1.157199
Denmark	1.000741	0.988889	0.997778	1.003704	0.972593†	1.103704•	1.045926
France	0.962085	0.954502	0.954976	0.965877	0.918009†	1.155924•	0.987204
Germany	0.942222	0.924815†	0.928889	0.942222	0.956667	1.122593•	0.968519
Hong Kong	1.174405	1.189881	1.197024	1.160714	1.067857†	1.507143•	1.128571
India	1.067897	1.104059	1.064207	1.064576	1.037638†	1.202952•	1.071587
Indonesia	1.248333	1.253333	1.248889	1.187778	1.247222	1.440556•	1.161667†
Ireland	1.066355	1.062617	1.074766	1.072897	1.009346†	1.156075•	1.080374
Israel	1.192763	1.2125	1.226316	1.204605	1.071711†	1.267763•	1.161184
Japan	1.081865	1.067358	1.082383	1.082383	1.015544†	1.127461•	1.05285
Korea	1.106105	1.107267	1.118314	1.110174	1.1	1.294767•	1.097384†
Malaysia	1.275776	1.270807	1.301863	1.231056	1.197516†	2.909317•	1.631056
Netherlands	0.953419	0.937179†	0.944444	0.952564	0.955556	1.187607•	1.000427
Philippines	1.31497	1.248503	1.279042	1.313772	1.252695	1.449102•	1.237126†
Singapore	1.17561	1.170732	1.187805	1.155285	1.122764†	1.458537•	1.144715
Spain	0.977778	0.968333	0.97	0.983333	0.906111†	1.188333•	1.000556
Sweden	0.996525	0.971815	0.99305	0.995753	0.962548†	1.137838•	1.001158
Thailand	1.136082	1.160825	1.139175	1.114948	1.064948†	1.455155•	1.16134
Turkey	1.119016	1.102128	1.12008	1.096809	1.006383†	1.301064•	1.120213
UK	0.990226	0.975188†	1.001504	0.988722	0.978947	1.174436•	1.026316
USA	1.012403	0.98062†	0.988372	1.012403	1.022481	1.134884•	1.005426

Notes: †: Best performing model in the row, •:Worst performing model in the row.

The Mean Absolute Percentage Error (MAPE) is used. This is given by:

$$MAPE = \frac{100}{T-(T_1-1)} \sum_{t=T_1}^{T} \left| \frac{y_{t+s} - f_{t,s}}{y_{t+s}} \right|$$

Where T is the total sample size (in-sample and out-of-sample), T_1 is the first out-of-sample forecast observation, y_{t+s} is the average actual forecast values and $f_{t,s}$ is the forecast values. The attractive property of the MAPE is that it can be interpreted as a percentage error, and furthermore its value is bounded from below by zero.

Appendix 4

$$\log(\sigma_t^2) = \alpha + \beta \log(h_t^2) + \varepsilon$$

Australia			
Model	β	p-stat	R^2
GARCH	1.696156	0.000	0.045568
EGARCH	1.597178	0.000	0.06859
TGARCH	1.737834	0.000	0.06231
CGARCH	1.54809	0.000	0.044473
HYGARCH	0.845222	0.000	0.046248
Exp.Sm	0.792082	0.017	0.003446
MA	1.093639	0.000	0.049323

Austria			
Model	β	p-stat	R^2
GARCH	1.119578	0.000	0.039898
EGARCH	1.00016	0.000	0.040703
TGARCH	1.044136	0.000	0.042064
CGARCH	0.964538	0.000	0.036949
HYGARCH	0.751205	0.000	0.028614
Exp.Sm	-0.3163	0.4069	0.000422
MA	1.276699	0.000	0.03689

Belgium			
Model	β	p-stat	R^2
GARCH	1.18596	0.000	0.135383
EGARCH	1.356127	0.000	0.141745
TGARCH	1.17213	0.000	0.141938
CGARCH	1.163812	0.000	0.134442
HYGARCH	1.133702	0.000	0.132346
Exp.Sm	0.408058	0.199	0.000986
MA	1.16868	0.000	0.126833

Brazil			
Model	β	p-stat	R^2
GARCH	2.969954	0.000	0.021821
EGARCH	0.789803	0.000	0.025038
TGARCH	2.908369	0.000	0.022215
CGARCH	2.54824	0.000	0.018177
HYGARCH	0.557872	0.000	0.010653
Exp.Sm	0.76021	0.000	0.016005
MA	1.563481	0.000	0.021134

Chile			
Model	β	p-stat	R^2
GARCH	0.930077	0.000	0.02948
EGARCH	0.845867	0.000	0.032036
TGARCH	0.936852	0.000	0.030421
CGARCH	0.855989	0.000	0.030172
HYGARCH	0.177039	0.106	0.001592
Exp.Sm	0.344676	0.214	0.000944
MA	1.219717	0.000	0.023074

Denmark			
Model	β	p-stat	R^2
GARCH	1.141096	0.000	0.075447
EGARCH	1.325983	0.000	0.081294
TGARCH	1.135601	0.000	0.075492
CGARCH	1.142596	0.000	0.078443
HYGARCH	1.013771	0.000	0.080206
Exp.Sm	0.332954	0.477	0.000308
MA	1.025941	0.000	0.063922

France			
Model	β	p-stat	R^2
GARCH	1.388381	0.000	0.108307
EGARCH	1.418669	0.000	0.128305
TGARCH	1.496607	0.000	0.121171
CGARCH	1.35E+00	0.000	0.10666
HYGARCH	1.05522	0.000	0.115212
Exp.Sm	0.043401	0.895	0.00001
MA	1.235262	0.000	0.115801

Germany			
Model	β	p-stat	R^2
GARCH	1.355733	0.000	0.168525
EGARCH	1.432755	0.000	0.180631
TGARCH	1.402096	0.000	0.178223
CGARCH	1.306784	0.000	0.167111
HYGARCH	1.282348	0.000	0.16705
Exp.Sm	-0.22744	0.399	0.000426
MA	1.265875	0.000	0.164468

Hong Kong			
Model	beta	p-stat	R^2
GARCH	1.243288	0.000	0.069553
EGARCH	1.146957	0.000	0.079649
TGARCH	1.216809	0.000	0.070948
CGARCH	1.252011	0.000	0.072943
HYGARCH	0.914235	0.000	0.075297
Exp.Sm	2.753356	0.000	0.080764
MA	1.318519	0.000	0.097326

India			
Model	β	p-stat	R^2
GARCH	0.947932	0.000	0.08643
EGARCH	0.95682	0.000	0.079971
TGARCH	0.95252	0.000	0.085216
CGARCH	0.920876	0.000	0.084099
HYGARCH	0.840273	0.000	0.08369
Exp.Sm	0.979743	0.018	0.003396
MA	0.965007	0.000	0.071761

Indonesia			
Model	β	p-stat	R^2
GARCH	0.428123	0.000	0.013849
EGARCH	0.409744	0.000	0.009814
TGARCH	0.429661	0.000	0.014002
CGARCH	0.364152	0.000	0.010773
HYGARCH	0.37385	0.000	0.012528
Exp.Sm	0.821907	0.010	0.004196
MA	0.325492	0.025	0.003127

Ireland			
Model	β	p-stat	R^2
GARCH	0.912537	0.000	0.05127
EGARCH	0.961149	0.000	0.057374
TGARCH	0.891837	0.000	0.053399
CGARCH	0.887287	0.000	0.049592
HYGARCH	0.851362	0.000	0.055762
Exp.Sm	1.747766	0.001	0.006538
MA	1.018956	0.000	0.05428

Israel			
Model	β	p-stat	R^2
GARCH	1.076547	0.000	0.03673
EGARCH	1.055811	0.000	0.038328
TGARCH	1.129789	0.000	0.039744
CGARCH	1.031699	0.000	0.036969
HYGARCH	0.753095	0.000	0.038114
Exp.Sm	1.923445	0.000	0.015835
MA	1.159075	0.000	0.029877

Japan			
Model	β	p-stat	R^2
GARCH	0.831533	0.000	0.029201
EGARCH	0.793476	0.000	0.04304
TGARCH	0.793069	0.000	0.038207
CGARCH	0.794501	0.000	0.027185
HYGARCH	0.615647	0.000	0.030569
Exp.Sm	2.551355	0.000	0.011652
MA	1.175483	0.000	0.043461

Korea			
Model	β	p-stat	R^2
GARCH	0.877843	0.000	0.060038
EGARCH	0.865994	0.000	0.062476
TGARCH	0.861182	0.000	0.064896
CGARCH	0.892698	0.000	0.061253
HYGARCH	0.849804	0.000	0.06259
Exp.Sm	2.761219	0.000	0.045301
MA	1.053787	0.000	0.072586

Malaysia			
Model	β	p-stat	R^2
GARCH	1.093134	0.000	0.088146
EGARCH	0.974076	0.000	0.094522
TGARCH	1.090729	0.000	0.087106
CGARCH	1.098223	0.000	0.096061
HYGARCH	0.838919	0.000	0.104146
Exp.Sm	1.939317	0.000	0.09359
MA	2.361407	0.000	0.104399

Netherlands			
Model	β	p-stat	R^2
GARCH	1.09416	0.000	0.166345
EGARCH	1.162861	0.000	0.17102
TGARCH	1.115016	0.000	0.173001
CGARCH	1.096981	0.000	0.16697
HYGARCH	1.069587	0.000	0.166737
Exp.Sm	0.223873	0.3197	0.000596
MA	1.083225	0.000	0.154545

Philippines			
Model	β	p-stat	R^2
GARCH	0.472796	0.000	0.015109
EGARCH	0.504206	0.000	0.014063
TGARCH	0.497386	0.000	0.015682
CGARCH	0.466225	0.000	0.014871
HYGARCH	0.459382	0.000	0.016184
Exp.Sm	0.101553	0.7317	0.000072
MA	0.503476	0.000	0.01026

Singapore			
Model	β	p-stat	R^2
GARCH	0.964814	0.000	0.074338
EGARCH	0.936179	0.000	0.082899
TGARCH	0.985312	0.000	0.079065
CGARCH	0.947702	0.000	0.075242
HYGARCH	0.806221	0.000	0.07814
Exp.Sm	2.539198	0.000	0.054346
MA	1.192595	0.000	0.088287

Spain			
Model	β	p-stat	R^2
GARCH	1.473623	0.000	0.143648
EGARCH	1.453446	0.000	0.162501
TGARCH	1.540177	0.000	0.153741
CGARCH	1.452799	0.000	0.140816
HYGARCH	0.958395	0.000	0.147465
Exp.Sm	2.812581	0.000	0.032292
MA	1.479299	0.000	0.155914

Sweden			
Model	β		R^2
GARCH	1.178774	0.000	0.127222
EGARCH	1.225548	0.000	0.139004
TGARCH	1.195745	0.000	0.136453
CGARCH	1.179363	0.000	0.125864
HYGARCH	0.938374	0.000	0.132502
Exp.Sm	1.625786	0.000	0.011254
MA	1.243511	0.000	0.137961

Thailand			
Model	β	p-stat	R^2
GARCH	0.96317	0.000	0.049911
EGARCH	0.884114	0.000	0.048868
TGARCH	0.931476	0.000	0.051094
CGARCH	0.960676	0.000	0.048948
HYGARCH	0.823965	0.000	0.049805
Exp.Sm	1.814162	0.000	0.027639
MA	1.346386	0.000	0.046588

Turkey			
Model	β	p-stat	R^2
GARCH	1.183826	0.000	0.075764
EGARCH	1.204875	0.000	0.075544
TGARCH	1.19119	0.000	0.076125
CGARCH	1.118928	0.000	0.084202
HYGARCH	0.797433	0.000	0.08476
Exp.Sm	2.408521	0.000	0.040475
MA	1.241275	0.000	0.083083

UK			
Model	β	p-stat	R^2
GARCH	1.063285	0.000	0.124262
EGARCH	1.106444	0.000	0.13399
TGARCH	1.068037	0.000	0.134412
CGARCH	1.063328	0.000	0.124269
HYGARCH	1.040288	0.000	0.124564
Exp.Sm	0.26102	0.425	0.000383
MA	1.020204	0.000	0.117053

USA			
Model	β	p-stat	R^2
GARCH	1.042832	0.000	0.096817
EGARCH	1.196328	0.000	0.108545
TGARCH	1.123456	0.000	0.109318
CGARCH	1.040516	0.000	0.097257
HYGARCH	1.128023	0.000	0.099692
Exp.Sm	1.45716	0.000	0.007936
MA	1.132793	0.000	0.104894

$$\sigma_t^2 = \alpha + \beta h_t^2 + \varepsilon_t$$

Australia			
Model	β	p-stat	R^2
GARCH	1.041815	0.000	0.04396
EGARCH	1.343178	0.000	0.066153
TGARCH	0.921163	0.000	0.053811
CGARCH	0.980779	0.000	0.039182
HYGARCH	0.988623	0.000	0.039671
Exp.Sm	0.930828	0.006	0.00445
MA	0.95765	0.000	0.036237

Austria			
Model	β	p-stat	R^2
GARCH	0.821496	0.000	0.06921
EGARCH	0.951315	0.000	0.075368
TGARCH	0.796154	0.000	0.074331
CGARCH	0.627671	0.000	0.056328
HYGARCH	0.594548	0.000	0.050794
Exp.Sm	-0.09307	0.763	0.000053
MA	0.990147	0.000	0.041613

Belgium			
Model	β	p-stat	R^2
GARCH	1.12959	0.000	0.171634
EGARCH	1.878555	0.000	0.208281
TGARCH	1.266061	0.000	0.196511
CGARCH	1.126673	0.000	0.182897
HYGARCH	1.170336	0.000	0.180039
Exp.Sm	0.450652	0.225	0.00086
MA	1.09542	0.000	0.114995

Brazil			
Model	β	p-stat	R^2
GARCH	0.886215	0.000	0.039787
EGARCH	0.23438	0.000	0.050604
TGARCH	0.864931	0.000	0.040794
CGARCH	0.80512	0.000	0.035528
HYGARCH	0.661711	0.000	0.023609
Exp.Sm	0.160381	0.000	0.013501
MA	0.974264	0.000	0.038632

Chile			
Model	β	p-stat	R^2
GARCH	0.636809	0.000	0.06318
EGARCH	0.63485	0.000	0.067711
TGARCH	0.634973	0.000	0.06446
CGARCH	0.605356	0.000	0.062715
HYGARCH	0.15315	0.000	0.007883
Exp.Sm	0.442356	0.000	0.007579
MA	1.004601	0.000	0.045731

Denmark			
Model	β	p-stat	R^2
GARCH	1.149203	0.000	0.120924
EGARCH	1.558682	0.000	0.139739
TGARCH	1.157789	0.000	0.129642
CGARCH	1.131261	0.000	0.11806
HYGARCH	1.196857	0.000	0.127772
Exp.Sm	0.038791	0.930	0.000005
MA	0.980625	0.000	0.075626

France			
Model	β	p-stat	R^2
GARCH	1.418071	0.000	0.193645
EGARCH	1.661976	0.000	0.223055
TGARCH	1.538335	0.000	0.214481
CGARCH	1.36E+00	0.000	0.188889
HYGARCH	1.362079	0.000	0.191132
Exp.Sm	0.755398	0.010	0.00391
MA	1.126771	0.000	0.159555

Germany			
Model	β	p-stat	R^2
GARCH	1.180864	0.000	0.207901
EGARCH	1.473375	0.000	0.236032
TGARCH	1.266293	0.000	0.229822
CGARCH	1.105785	0.000	0.219485
HYGARCH	1.159607	0.000	0.21119
Exp.Sm	0.590375	0.020	0.003169
MA	1.079327	0.000	0.182333

Hong Kong			
Model	β	p-stat	R^2
GARCH	0.727049	0.000	0.050854
EGARCH	0.813386	0.000	0.071098
TGARCH	0.631489	0.000	0.053898
CGARCH	0.739149	0.000	0.048178
HYGARCH	0.726178	0.000	0.049705
Exp.Sm	1.415401	0.000	0.049282
MA	0.998872	0.000	0.075205

India			
Model	β	p-stat	R^2
GARCH	0.837414	0.000	0.116757
EGARCH	0.822921	0.000	0.096908
TGARCH	0.853095	0.000	0.113226
CGARCH	0.879364	0.000	0.132049
HYGARCH	0.867701	0.000	0.125522
Exp.Sm	1.074607	0.010	0.003822
MA	0.954494	0.000	0.077318

Indonesia			
Model	β	p-stat	R^2
GARCH	0.366267	0.000	0.019721
EGARCH	0.456354	0.000	0.019153
TGARCH	0.365994	0.000	0.019941
CGARCH	0.258588	0.000	0.013121
HYGARCH	0.24686	0.000	0.015596
Exp.Sm	0.386347	0.103	0.001547
MA	0.775465	0.000	0.015944

Ireland			
Model	β	p-stat	R^2
GARCH	0.882305	0.000	0.059343
EGARCH	0.991538	0.000	0.071966
TGARCH	0.860066	0.000	0.064593
CGARCH	0.849098	0.000	0.057053
HYGARCH	0.961744	0.000	0.069598
Exp.Sm	0.877276	0.091	0.001665
MA	0.993518	0.000	0.049063

Israel			
Model	β	p-stat	R^2
GARCH	0.698877	0.000	0.037987
EGARCH	0.824397	0.000	0.048822
TGARCH	0.600901	0.000	0.039416
CGARCH	0.655596	0.000	0.04095
HYGARCH	0.639499	0.000	0.040147
Exp.Sm	1.560657	0.000	0.014729
MA	1.034905	0.000	0.035839

Japan			
Model	β	p-stat	R^2
GARCH	0.714297	0.000	0.041922
EGARCH	0.717114	0.000	0.04886
TGARCH	0.601667	0.000	0.043274
CGARCH	0.689388	0.000	0.039138
HYGARCH	0.688898	0.000	0.040533
Exp.Sm	1.715778	0.000	0.008291
MA	0.991035	0.000	0.042239

Korea			
Model	β		R^2
GARCH	0.66315	0.000	0.046965
EGARCH	0.744956	0.000	0.051146
TGARCH	0.619496	0.000	0.050049
CGARCH	0.711129	0.000	0.052929
HYGARCH	0.686132	0.000	0.051896
Exp.Sm	2.005937	0.000	0.034312
MA	1.033731	0.000	0.071023

Malaysia			
Model	β	p-stat	R^2
GARCH	0.642556	0.000	0.061115
EGARCH	0.673722	0.000	0.059108
TGARCH	0.567945	0.000	0.057639
CGARCH	0.679057	0.000	0.068913
HYGARCH	0.673043	0.000	0.091248
Exp.Sm	0.581774	0.000	0.053951
MA	1.199087	0.000	0.063098

Netherlands			
Model	β	p-stat	R^2
GARCH	1.077871	0.0000	0.232284
EGARCH	1.32963	0.0000	0.253808
TGARCH	1.133367	0.0000	0.251496
CGARCH	1.070949	0.0000	0.229004
HYGARCH	1.125014	0.0000	0.233859
Exp.Sm	0.630169	0.0128	0.00361
MA	0.990671	0.0000	0.166529

Philippines			
Model	β	p-stat	R^2
GARCH	0.16536	0.0583	0.00209
EGARCH	0.413876	0.0057	0.004452
TGARCH	0.289747	0.0086	0.00402
CGARCH	0.178411	0.0367	0.002543
HYGARCH	0.159997	0.0595	0.00207
Exp.Sm	0.873769	0.00735	0.001868
MA	1.382132	0.000	0.031526

Singapore			
Model	β	p-stat	R^2
GARCH	0.520196	0.000	0.037921
EGARCH	0.613579	0.000	0.047449
TGARCH	0.451003	0.000	0.037497
CGARCH	0.470591	0.000	0.029824
HYGARCH	0.487662	0.000	0.033672
Exp.Sm	1.559816	0.000	0.029458
MA	0.973682	0.000	0.04862

Spain			
Model	β	p-stat	R^2
GARCH	1.190015	0.000	0.173053
EGARCH	1.335723	0.000	0.199308
TGARCH	1.271292	0.000	0.191503
CGARCH	1.146993	0.000	0.163866
HYGARCH	1.181418	0.000	0.174551
Exp.Sm	1.719337	0.000	0.020883
MA	1.125969	0.000	0.146667

Sweden			
Model	β	p-stat	R^2
GARCH	0.916065	0.000	0.115183
EGARCH	1.155834	0.000	0.146058
TGARCH	0.963765	0.000	0.140443
CGARCH	0.914544	0.000	0.113985
HYGARCH	0.914658	0.000	0.115663
Exp.Sm	0.714136	0.0295	0.002763
MA	1.060244	0.000	0.110553

Thailand			
Model	β	p-stat	R^2
GARCH	0.718725	0.000	0.083539
EGARCH	0.668834	0.000	0.077504
TGARCH	0.673039	0.000	0.093475
CGARCH	0.828942	0.000	0.098288
HYGARCH	0.810749	0.000	0.102129
Exp.Sm	1.108245	0.000	0.033662
MA	1.066714	0.000	0.059608

Turkey			
Model	β	p-stat	R^2
GARCH	0.930548	0.000	0.116589
EGARCH	1.300413	0.000	0.11279
TGARCH	0.938564	0.000	0.120133
CGARCH	0.980769	0.000	0.130186
HYGARCH	0.950526	0.000	0.12557
Exp.Sm	1.944308	0.000	0.024074
MA	1.262409	0.000	0.092183

UK			
Model	β	p-stat	R^2
GARCH	1.086905	0.000	0.192248
EGARCH	1.292572	0.000	0.212622
TGARCH	1.080371	0.000	0.210859
CGARCH	1.088978	0.000	0.195262
HYGARCH	1.16052	0.000	0.20255
Exp.Sm	0.700018	0.026	0.002904
MA	0.967601	0.000	0.140412

USA			
Model	β	p-stat	R^2
GARCH	0.966411	0.000	0.125702
EGARCH	1.287761	0.000	0.180725
TGARCH	1.116444	0.000	0.169945
CGARCH	0.944967	0.000	0.123778
HYGARCH	1.048889	0.000	0.134925
Exp.Sm	0.805075	0.0209	0.003107
MA	1.068156	0.000	0.125414

Chapter 9

Appendix 5

Are RiskMetrics forecasts good enough? Evidence from 31 stock markets

David G. McMillan, Dimos Kambouroudis
School of Management, University of St Andrews, UK

International Review of Financial Analysis 18 (2009) 117–124
doi:10.1016/j.irfa.2009.03.006