ASSET PRICE AND VOLATILITY FORECASTING USING NEWS SENTIMENT

ZRYAN A SADIK

A thesis submitted for the degree of Doctor of Philosophy

Department of Mathematics
College of Engineering, Design and Physical Sciences

Brunel University London

October 2018

Dedication

This thesis is dedicated to my parents . . .

Abstract

The aim of this thesis is to show that news analytics data can be utilised to improve the predictive ability of existing models that have useful roles in a variety of financial applications. The modified models are computationally efficient and perform far better than the existing ones. The new modified models offer a reasonable compromise between increased model complexity and prediction accuracy.

I have investigated the impact of news sentiment on volatility of stock returns. The GARCH model is one of the most common models used for predicting asset price volatility from the return time series. In this research, I have considered quantified news sentiment as a second source of information and its impact on the movement of asset prices, which is used together with the asset time series data to predict the volatility of asset price returns. Comprehensive numerical experiments demonstrate that the new proposed volatility models provide superior prediction than the "plain vanilla" GARCH, TGARCH and EGARCH models. This research presents evidence that including news sentiment term as an exogenous variable in the GARCH framework improves the prediction power of the model. The analysis of this study suggested that the use of an exponential decay function is good when the news flow is frequent, whereas the Hill decay function is good only when there are scheduled announcements. The numerical results vindicate some recent findings regarding the utility of news sentiment as a predictor of volatility, and also vindicate the utility of the new models combining the proxies for past news sentiments and the past asset price returns. The empirical analysis suggested that news augmented GARCH models can be very useful in estimating VaR and implementing risk management strategies.

Another direction of my research is introducing a new approach to construct a commodity futures pricing model. This study proposed a new method of incorporating macroeconomic news into a predictive model for forecasting prices of crude oil futures contracts. Since these futures contracts are more liquid than the underlying commodity itself, accurate forecasting of their prices is of great value to multiple categories of market participants. The Kalman filtering framework for forecasting arbitrage-free (futures) prices was utilized, and it is assumed that the volatility of oil (futures) price is influenced by macroeconomic news. The impact of quantified news sentiment on the price volatility is modelled through a parametrized, non-linear functional map. This approach is motivated by the successful use of a similar model structure in my earlier work, for predicting individual stock volatility using stock-specific news. Numerical experiments with real data illustrate that this new model performs better than the one factor model in terms of accuracy of predictive power as well as goodness of fit to the data. The proposed model structure for incorporating macroeconomic news together with historical (market) data is novel and improves the accuracy of price prediction quite significantly.

Acknowledgements

Completion of this doctoral thesis was possible with the support of several people, who have so generously contributed in different ways to the work presented in this thesis. I would like to express my sincere gratitude to all of them.

Foremost, I would like to express my special thanks and deep appreciation to my supervisor Dr Paresh Date for the continuous support of my PhD study and related research, for his immense knowledge, patience, motivation, commitment and constructive feedback that followed every step of my research. His valuable guidance helped me in all the time of my research and writing of this thesis. I would also like to thank him for helping me to grow as a research scientist. I could not have imagined having a better supervisor and mentor for my PhD study.

Besides my supervisor, my special gratitude goes to my industrial supervisor Professor Gautam Mitra from OptiRisk Systems for his insightful comments, constructive suggestions and his enthusiastic support that enabled me to complete my thesis. His valuable guidance and advice have been absolutely vital for the development of this research.

This thesis would not have come to a successful completion, without the help I received from the staff members of OptiRisk Systems. In particular, my present and past colleagues, Dr Christina Erlwein-Sayer, Dr Xiang Yu, Dr Cristiano Arbex Valle, Dr Tilman Sayer, Dr Christian Valente and Mr Ansuman Swain. They have all extended their support in a very special way. I gained a lot from them through their personal and scholarly interactions, and their suggestions at various points of my research programme. I am extremely grateful to all of them especially Dr Christina for her constructive criticism and proofreading my thesis.

I would also like to extend my sincere gratitude to OptiRisk Systems. I am very grateful for their financial support and also for providing me the experimental data. Without their precious support it would not be possible to conduct this research.

I gratefully acknowledge the academic staff of the Department of Mathematics at Brunel University for their help. I am particularly thankful to Dr Cormac Lucas, Dr Elena Boguslavskaya, Dr Anne-Sophie Kaloghiros and Dr David Meier for giving me the opportunity to attend their MSc classes, which were related to financial mathematics, in the first year of my PhD study. I am also very thankful to Dr Veronica Vinciotti for providing me her lecture notes when I first started my PhD study. I am also grateful to the staff members of the department, especially Ms Frances Foster, Mr Nalin Soni and Ms Deborah Banyard for all their assistance.

Finally, but by no means least, I am deeply thankful to my family: my parents and my sibling for their unconditional love and unbelievable support. I owe a lot to my parents, who encouraged and helped me at every stage of my personal and academic life, and longed to see this achievement come true. They are the most important people in my world and I dedicate this thesis to them.

Contents

List of Abbreviations

ARCH	Autoregressive Conditional Heterscedasticity
GARCH	Generalized Autoregressive Conditional Heterscedasticity
GARCH-t	GARCH with t-distributed residuals
EGARCH	Exponential GARCH
TGARCH	Threshold GARCH
PGARCH	Periodic GARCH
CGARCH	Component GARCH
AGARCH	Asymmetric GARCH
GARCH-JUMP	GARCH with jump
FIGARCH	Fractionally Integrated GARCH
ATGARCH	Asymmetric Threshold GARCH
NA1-GARCH	First News Augmented GARCH
NA2-GARCH	Second News Augmented GARCH
GED-GARCH	Generalized Error Distribution GARCH
GRS-GARCH	Generalized Regime-Switching GARCH
GJR-GARCH	Glosten, Jagannathan and Runkle GARCH
SEMIFAR-GARCH	Semiparametric Fractional Autoregressive GARCH
VAR	Vector Autoregressive Model
VaR	Value-at-Risk
LSE	Least Squares Estimation
MLE	Maximum Likelihood Estimation
ESS	Event Sentiment Score
ENS	Event Novelty Score
AIC	Akaike's Information Criterion
AICc	Akaike's Information Criterion (Second Order)
BIC	Bayesian information criterion
MAE	Mean Absolute Error
RMSE	Root Mean Squared Error
MAPE	Mean Average Percentage Error

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Research Background and Context

In the domain of finance, characterization and quantification of uncertainty plays a crucial role, both in risk management and portfolio optimization. Volatility of the asset returns is one of the most common proxies for uncertainty, which cannot be observed directly. Improving the accuracy of volatility estimates and volatility forecasts can be hugely beneficial for better risk management. Hedge funds use volatility forecasting to design their asset allocation and hedging strategies. Volatility estimates are also important in portfolio optimization and in pricing of financial derivatives. Market participants (market makers, banks, option traders) as well as market regulators need tools for forecasting volatility of prices of various financial instruments accurately.

In the last few decades, forecasting and the analysis of asset returns' volatility has attracted a lot of attention in academic and industrial research. The most straightforward approach to estimate (squared) volatility is to find the sample conditional variance of the underlying asset price returns from historic time series data. However, this ignores several stylised features of the time series of asset price returns, e.g. volatility clustering, where large changes in the asset returns tend to be followed by large changes and vice versa. A class of models, which we refer broadly as Generalised autoregressive heteroskedastisity (GARCH) type models, express volatility at a point in time as a function of past volatility and past asset price returns. Models from this class tend to forecast volatility far more accurately that simple conditional sample variance and are employed widely in the finance industry for pricing and risk management purposes.

GARCH type models still suffer from one major shortcoming: the sources of impact on volatility are solely past asset price returns and past volatilities. In practice, volatility of asset returns is also affected by investor sentiment, which in turn is affected by news related to that asset. The news can be time-synchronous (e.g. scheduled earnings report for a company) or time-asynchronous (e.g. an industrial accident or an unexpected election result). Arrival of these two types of information impacts the stock market at the time of release and persists over finite periods of time, from minutes to days. In recent years, commercial data analytics providers have started providing quantitative measures called *news sentiment scores* for individual stocks as well as for a variety of macroeconomic news items. The news analytics providers collect and interpret news articles from an extensive range of worldwide sources. In practice, news analytics refers to the automated process of computational linguistic analysis to measure quantitative and qualitative characteristics of news stories. This process employs text and data mining tools and techniques to provide summary information by extracting action phrases and important keywords from the incoming news articles. The process allocates a score within a specified range for each asset (e.g. a stock or a commodity) specific news item, which indicates the extent to which the investor sentiment is positive or negative for that asset at the specified point in time. Typically, each news item is given attributes such as relevance, novelty and sentiment scores according to an individual entity. The analytical process of producing such scores is fully automated from collecting, extracting, aggregating, to categorising and scoring. The output is an individual score assigned to each news article for each attribute. A *sentiment score* of a news item measures the emotional tone within the text and varies between positive, neutral and negative. Companies such as Ravenpack and Thomson Reuters provide data analytics on a wide range of stocks, commodities and fixed income instruments.

Sentiment analysis has greatly attracted a large amount of research in recent years, especially after the speedy spread of social media networks, forums and blogs. The effortlessness of approachability to news analytics and other information sources have become useful components for analysts and market participants to enhance their models and trading strategies in the financial markets. With all the disseminating news on the news vendors' and brokers' platforms, the trading opportunities for making profits have become endless. Based on such sufficient information making a trader judgement and reaching at a right decision is only a matter of how long time it will take. Hence, the automated procedure of analysing news sentiment metadata has helped significantly to improve the process of making decisions in terms of time and efficiency.

The development of a new model or modifying an existing model for estimating and forecasting asset price and volatility have always had a certain appeal to researchers in the finance applications, particularly concerning the risk management theory which is based on that statistical measurement. To banking institutions and market makers in general, these new or modified approaches should be robust and better comprehensible to forecast, measure and estimate the variable of interest. The current approaches, like conditional variance or stochastic volatility, have not incorporated the impact of news into volatility prediction models in a way which is meaningful from an economic point of view. To researchers in academia and business, it is important to develop a direct way of measuring volatility, which cannot be observed directly, by implementing and modifying such models that can consistently outperform predictions and capture all stylized facts of financial time series.

To apply news sentiment data effectively and efficiently in trading decisions and risk management, one has to be able to identify news items which are both relevant and new (currently reported). In other words, it is important that the researchers have the ability of identifying related information events and distinguishing "old" news items (previously reported) from genuinely "new" news items. This is especially true for those applications that use high frequency data, since their algorithms have to make a response swiftly to new information. Finally, to link daily asset closing prices to news sentiment data, the news sentiment scores have to be aggregated into daily numbers. This aggregation of sentiment scores over a day will help reducing or eliminating the noise in experimental data.

There is a small but growing academic literature that links sentiment and news flows (intensity) into efficient volatility forecasting for equities. In order to use this predictive power effectively for analysing and estimating a portfolio's risk, the prediction of risk has to be structured as a function of all assets' volatilities that formed the portfolio. To date, the influence of exogenous variables such as firm-specific and macroeconomic news sentiment on several financial time series in the stock and commodity markets has not been sufficiently explored in the academic literature.

The objective of this thesis is to investigate, evaluate and identify valuable features of news and to show improvement of the predictive power of the existing asset price and volatility models by including a proxy of news sentiment as an exogenous variable in the model structure. Therefore, the overall aim is to explore the use of news sentiment scores to improve financial forecasting. We will look at forecasting of volatility, risk forecasting and commodity price forecasting, aided in all the cases by time series of corresponding news sentiment scores. We also explore, to some extent, the relationship between stock markets, news sentiment data and macroeconomic indices. We then investigate the impact of news on financial variables and construct a number of predictive models to estimate and predict asset volatility as well as the crude oil futures prices utilising news sentiment data.

This thesis looks at asset price and volatility forecasting models using news sentiment data. The contribution of it lies in (i) proposing a method of incorporating firm-specific news information into a predictive model for estimating and predicting the volatility of stock/asset price returns for a certain number of stock indices and assets. The quantified news sentiment and its impact on the movement of asset prices is considered as a second source of information, which is used together with the asset price data to predict the volatility of asset price returns. (ii) Proposing a method of incorporating macroeconomic news into a predictive model for forecasting prices of crude oil futures contracts. The impact of quantified news sentiment on the price volatility is modelled through a functional map. Further discussion on commodity and futures prices modelling will be presented in chapter 5.

In this thesis, the proposed model structures for incorporating news sentiment data together with historical (market) data are novel and significant in accurate prediction. The choice of the model structures are essentially heuristic and justified through numerical experiments. Finally, the proposed models are also more powerful than the existing benchmark ones in terms of out-of-sample performance (predictive power), and yield a comparable accuracy.

1.2 Organization of Thesis

The rest of the chapters are organised as follows:

- Chapter 2 outlines the necessary mathematical and financial preliminary information. The first section is focused on the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) type models. A brief description and overview of a number of symmetric and asymmetric GARCH models is provided. The next section describes common methods for estimation of model parameters for this class of models. The last section provides preliminary information of the Value-at-Risk (VaR), its calculation and model backtesting used in assessing the validity of VaR models.

- Chapter 3 discusses the development of a predictive volatility model using news sentiment data. The idea is to utilize the news sentiment data as a second source of information to enhance stock returns volatility predictions by combining the information obtained from symmetric GARCH modelling and news sentiment scores. A practical measure (news impact score) has been used, which quantifies how a collection of news items impact volatility. This measure takes into consideration news flow and exponential decay of news sentiment over time.

This study looks at various different models from the family of modified GARCH models (such as EGARCH and TGARCH), and at various possible model structures to incorporate news sentiment scores. Particular emphasis has been given to the forecasting performance of the suggested models and whether they can capture the characteristics of asset volatility and the influence of news on it. A new modified version of the GARCH model has been proposed, which is called *first* news augmented GARCH (NA1-GARCH) model. The new model incorporates the impact of news into a volatility prediction model in a way which is meaningful from an economic point of view. Through numerical experiments on a large number of datasets from two different markets, this new model is shown to outperform several existing models in terms of predictive ability. Furthermore, this study carries out an empirical application to determine the VaR with different volatility estimates obtained by the aforementioned models.

- Chapter 4 investigates the impact of high-frequency public news sentiment on the daily log returns volatility. It considers quantified news sentiment and its impact on the movement of asset prices as a proxy for news sentiment, which is used together with the asset time series data to predict the volatility of asset price returns. This proxy is then used as an exogenous term in a modified version of a GARCH model, which is called *second* news augmented GARCH (NA2-GARCH) model. The model structure is novel and reflects the following economic realities: (i) positive and negative news impact the volatility differently,(ii) positive news tends to reduce volatility whereas negative news tends to increase volatility. The model structure of the NA2-GARCH model is also chosen to be one with a direct multiplicative effect of news on the GARCH-predicted volatility and this quantified effect of news is restricted by a lower and an upper bound to keep the news impact related scaling in a reasonable range. An extensive empirical investigation is carried out, on historical market data for 12 stocks from 2 different markets, to establish the superior predictive ability of this model.
- Chapter 5 focuses on the forecasting performance of different models that utilise the global macroeconomic news data as a proxy of news sentiment and its impact on the movement of crude oil prices. This study uses a one factor model with a

constant risk premium, a random mean and a seasonality adjustment in terms of an additive sinusoid. It also employs a linear state space model with logarithm of the spot price as the latent variable and a vector of logarithm of the futures prices as the observed variable. We analyse 12 futures contracts with different term maturities. In this study, two groups of models have been defined and then compared respectively to each other in terms of the predictive power. The two groups are one factor models and one factor models with macroeconomic news data.

A new method of incorporating macroeconomic news into a predictive model for forecasting prices of crude oil futures contracts has been proposed. This study has utilized the Kalman filtering framework for forecasting arbitrage-free (futures) prices, and assumes that the volatility of oil (futures) price is influenced by macroeconomic news. The impact of quantified news sentiment on the price volatility is modelled through a parametrized, non-linear functional map. This approach is motivated by the successful use of a similar model structure in Chapter 4, for predicting individual stock volatility using stock-specific news. The proposed model structure for incorporating macroeconomic news together with historical (market) data is novel. This chapter combines theory and practice by conducting a range of numerical experiments across the discussed models.

- Chapter 6 provides a conclusion of this thesis, along with future research directions.
- Appendix A and B provide the tables regarding the numerical experiments of the Chapters 3 and 4, respectively.
- Appendix C provides the tables of crude oil export and import countries of Chapter 5.
- Appendix D provides an overview of software developed for this research.

1.3 Publications

The contributory material from two chapters of this dissertation have been submitted to peer reviewed journals in terms of two papers:

- Chapter 4: "News Augmented GARCH(1,1) Model for Volatility Prediction" has been submitted to IMA Journal of Management Mathematics. **Accepted** for publication on 21 Feb 2018. Available online: https://academic.oup.com/imaman/advance-article/doi/10.1093/imaman/dpy004/4924966.

- Chapter 5: "Forecasting Crude Oil Futures Prices Using Global Macroeconomic News Sentiment" has been submitted to IMA Journal of Management Mathematics.

Chapter 2

Preliminaries

In the subsequent chapters, we will look at the use of news-enhanced versions of a certain class of volatility models and volatility forecasting and risk prediction. We will look at the underlying class of models in the next section.

2.1 Models From GARCH family

2.1.1 ARCH Model

The first and simplest model we will look at is an ARCH model, which stands for Autoregressive Conditional Heteroscedasticity. The model was proposed by Engle (1982). The AR comes from the fact that these models are autoregressive models in squared return. The conditional comes from the fact that in these models, next period's volatility is conditional on information of the current period. Heteroscedasticity means non constant volatility. In a standard linear regression where

$$y_i = \alpha + \beta x_i + \varepsilon_i \tag{2.1}$$

when the variance of the residuals, ε_i is constant, we call that homoscedastic and use ordinary least squares to estimate α and β. If, on the other hand, the variance of the residuals is not constant, we call that heteroscedastic and we can use weighted least squares to estimate the regression coefficients.

Let us assume that the return on an asset is

$$r_t = \mu + \sigma_t z_t \tag{2.2}$$

where z_t is a sequence of $N(0,1)$ i.i.d. random variables. We will define the residual return at time t, $r_t - \mu$, as

$$\varepsilon_t = \sigma_t z_t. \tag{2.3}$$

The simplest ARCH(1) model can be written as follows:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 \tag{2.4}$$

where $\alpha_0 > 0$ and $\alpha_1 \geq 0$ to ensure positive variance and $\alpha_1 < 1$ for stationarity. Under an ARCH(1) model, if the residual return, ε_t is large in magnitude, our forecast for next period's conditional volatility, σ_{t+1} will be large. We say that in this model, the returns are conditionally normal (conditional on all information up to time $t-1$, the one period returns are normally distributed).

2.1.2 GARCH Model

Until early 1980s, numerous models of prediction based on autoregression were put forward. In two landmark papers, Engle (1982) and Bollerslev (1986), the ARCH and GARCH (Generalized Autoregressive Conditional Heterscedasticity) models have been proposed respectively, and they are the most successful and popular models in predicting volatility. Their incredible popularity stems from their ability to capture, with a very flexible structure, some of the typical stylized facts of financial time series, such as *volatility clustering*, that is the tendency for volatility periods of similar magnitude -of either sign- to cluster; *leverage effect*, where changes in stock prices tend to be negatively correlated with changes in volatility, and *Long-range dependence* in the data, where sample autocorrelations of the data are small whereas the sample autocorrelations of the absolute and squared values are significantly different from zero even for large lags. Usually GARCH models can take into account the time-varying volatility phenomenon over a long period (see, for example, French et al. (1987) and Franses and Van Dijk (1996)) and provide very good in-sample estimates.

In an ARCH(1) model, the next period's variance only depends on the last period's squared residual so a crisis that caused a large residual would not have the sort of persistence that we observe after an actual crisis. This has led to an extension of the ARCH model to a GARCH(p,q), or Generalized ARCH model, that Bollerslev (1986) and Taylor (1986) developed independently of each other. The model allows the conditional variance of a variable to be dependent upon previous lags; first lag of the squared residual from the mean equation and present news about the volatility from the previous period which is as follows:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2 \ . \tag{2.5}$$

One of the most used and the simple model is the GARCH(1,1) process, for which the conditional variance is represented as a linear function of its own lags. The simplest GARCH (1,1) model can be written as follows:

$$\texttt{Mean equation} \qquad r_t = \mu + \varepsilon_t, \tag{2.6}$$

$$\texttt{Variance equation} \qquad \sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \tag{2.7}$$

where $\omega > 0,\ \alpha_1 \geq 0,\ \beta_1 \geq 0$, and
r_t = return of an asset at time t,
μ = average return,
ε_t = residual returns, defined as:

$$\varepsilon_t = \sigma_t z_t, \tag{2.8}$$

where z_t are standardized residual returns (i.e. i.i.d random variable with zero mean and unit variance), and σ_t^2 is conditional variance. If $\alpha_1 + \beta_1 < 1$ the model is covariance stationary and the unconditional variance equals $\sigma^2 = \omega/(1 - \alpha_1 - \beta_1)$.

For this unconditional variance to exist and be positive, we require that $\omega > 0, \alpha_1 > 0, \beta_1 > 0$. We impose these constrains so that our next period forecast of variance is a blend of our last period forecast and last period's squared return.

The GARCH model that has been described is typically called the GARCH(1,1) model. The (1,1) in parentheses is a standard notation in which the first number refers to how many autoregressive lags, or ARCH terms, appear in the equation, while the second number refers to how many moving average lags are specified, which here is often called the number of GARCH terms. Sometimes models with more than one lag are needed to find good variance forecasts.

2.1.3 GARCH-t Model

As we mentioned in the previous section, in 1986 an important contribution to this literature occurred when Bollerslev proposed the Generalized ARCH (GARCH) model as a more parsimonious way to capture volatility dynamics. In order to better account for the observed leptokurtosis, Bollerslev (1987) extended the GARCH specification to allow for the conditional Student's t-distribution as an alternative to the Normal. He argued that this formulation would permit us to distinguish between conditional leptokurtosis and conditional heteroskedasticity as plausible causes of the unconditional kurtosis observed in the data.

Financial modelling based on the Student's t-distribution, however, introduces an additional parameter (usually called *degrees of freedom*), which measures the extent of leptokurtosis in the data. One can also interpret this as a measure of the extent of departure from the Normal distribution. This in turn raises an estimation issue, since Zellner (1976) shows that maximum likelihood estimates do not exist for the linear regression coefficients, the dispersion parameter and the degrees of freedom parameter. Consequently, to use maximum likelihood, it is necessary to assign a degrees of freedom parameter that reflects the distributional properties of the error term. One commonly proposed technique for selecting the degrees of freedom is by using the kurtosis coefficient as a guide to solve for the implied degrees of freedom.

GARCH-t model can be specified in terms of its first two conditional moments. Specifically, the equation for r_t is given as follows:

$$r_t = \mu + \varepsilon_t \qquad \varepsilon_t / F_{t-1} \sim St_\nu(0, \sigma_t^2) \tag{2.9}$$

where μ is a constant and F_{t-1} denotes the σ-field generated by all the available information up through $t-1$, and ν is the degrees of freedom. The GARCH (p,q) conditional variance, σ_t^2 for this model takes the form:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \qquad p \geq 1, \qquad q \geq 1 \tag{2.10}$$

where the parameter restrictions $\omega > 0, \alpha_i \geq 0, \beta \geq 0$ ensure that the conditional variance is always positive. Moreover, $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$ is required for the convergence of the conditional variance. The distribution of the error term according to Bollerslev takes the form:

$$f(\varepsilon_t|Y_{t-1}^p;\theta_1) = \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\sigma_t^2(\nu-2)}}\left(1+\frac{\varepsilon_t^2}{\sigma_t^2(\nu-2)}\right)^{-(\nu+1)/2} \tag{2.11}$$

.

2.1.4 EGARCH Model

In his seminal paper, Nelson (1991) pointed out some major drawbacks of the GARCH theory indicating that a new approach is necessary to be established. The Exponential GARCH (EGARCH) model of Nelson meet these limitations. This model accounts for the fact that the volatility tends to rise in response to a bad news and fall in response to a good news, by introducing a sign term in the model. Further, the logarithm of the conditional variance is modelled, which means that no positivity condition needs to be imposed on the parameters to guarantee positivity of conditional variance. Nelson also derived a necessary and sufficient condition for strict stationarity of the EGARCH process, when ln σ_t^2 has an infinite moving average representation. In order to capture the asymmetric effect and in response to the other drawbacks of the GARCH models, in Nelson (1991), the exponential GARCH (EGARCH) model was proposed. EGARCH(p,q) model is represented as follows:

$$\ln(\sigma_t^2) = \omega + \sum_{i=1}^{p}(\alpha_i|z_{t-i}| + \gamma_i z_{t-i}) + \sum_{j=1}^{q}\beta_j \ln(\sigma_{t-i}^2), \tag{2.12}$$

where $z_{t-i} = \frac{\varepsilon_{t-i}}{\sigma_{t-i}}$ for $i = 1,\dots,q$ are standardized innovations. The parameter γ_i permits asymmetric effect and if $\gamma_i = 0$ then good news ($\varepsilon_{t-i} > 0$) will have the same effect on volatility as the bad news ($\varepsilon_{t-i} < 0$). The leverage effect can be examined by testing the assumption that $\gamma_i < 0$ as negative shocks will have bigger effects on volatility than positive shocks of the same magnitude. The log function is used in EGARCH models to ensure that the process remains positive. If shocks to variance $\ln(\sigma_t^2)$ perish fast and the deterministic is removed, then $\ln(\sigma_t^2)$ is strictly stationary as shown in Nelson (1991). The formula for the simple and popular process EGARCH(1,1) is given as follows:

$$\ln(\sigma_t^2) = \omega + \alpha_1|z_{t-1}| + \gamma_1 z_{t-1} + \beta_1 \ln(\sigma_{t-1}^2). \tag{2.13}$$

2.1.5 TGARCH Model

The Threshold GARCH (TGARCH) model is another modification of the ARCH model of Engle (1982) proposed by Zakoian (1994). In this model the conditional standard deviation is expressed as a piecewise linear function of the past white noise. This permits different response of the volatility to different signs of the lagged errors. The TGARCH model is given by

$$\varepsilon_t = \sigma_t z_t \tag{2.14}$$

$$\sigma_t = \omega + \sum_{i=1}^{q}(\alpha_i^+ \varepsilon_{t-i}^+ + \alpha_i^- \varepsilon_{t-i}^-) + \sum_{j=1}^{p} \beta_j \sigma_{t-j}, \tag{2.15}$$

where $z_t \sim N(0,1)$ independent of ε_{t-1} for all t, $\omega > 0,\ \alpha_i^+ \geq 0,\ \alpha_i^- \geq 0$ and $\beta_j \geq 0$ for all the values of i. $\varepsilon_t^+ = \max(\varepsilon_t, 0)$ and $\varepsilon_t^- = \min(\varepsilon_t, 0)$ are the positive and negative parts of ε_t and $(\alpha_i^+)_{i=1,\ldots,q}$, $(\alpha_i^-)_{i=1,\ldots,q}$ and $(\beta_j)_{j=1,\ldots,p}$ are real scalar sequences as described in Zakoian (1994). The effect of a shock ε_{t-i} on σ_t is a function of both its magnitude and its sign, and hence the model permits a response of volatility to news with different coefficients for good and bad news.

The simplest but very useful form of the model is TGARCH (1, 1) and is given as

$$\sigma_t = \omega + \alpha_1^+ \varepsilon_{t-1}^+ + \alpha_1^- \varepsilon_{t-1}^- + \beta_1 \sigma_{t-1}, \tag{2.16}$$

where $\omega > 0,\ \alpha_1^+ \geq 0$ and $\alpha_1^- \geq 0$.

A threshold GARCH model is similar to the EGARCH model as they both capture asymmetries in volatility but, they have some significant differences. Firstly the TGARCH makes volatility a function of (non-normalized) innovations, which the latter does not and hence it is closer to the classical formulations. Secondly in order to capture asymmetric effect EGARCH imposes a constant structure at all lags, while in the Threshold GARCH different lags may yield opposite contributions as explained in Zakoian (1994). TGARCH also preserves the stationary property and generates data with the fat tailed distribution as standard GARCH model does.

2.2 Parameter Estimation

2.2.1 Least Squares Estimation

The method of Least Squares Estimation (LSE) was first published by Adrain in 1808, but it was developed independently by Gauss in 1795, Legendre in 1805 and Adrain in 1808. The method is described as an algebraic procedure for fitting linear equations to data. The technique is about estimating model parameters by minimizing the sum of squared errors between observed data and expected values, or equivalently the sample average of squared errors. The method is a good choice when one is interested in finding the regression function that minimizes the corresponding expected squared error. In the least squares method the model parameters to be estimated must arise in expressions for the means of the observations. When the parameters appear linearly in these expressions, the problem of least squares estimation can be solved in closed form and it is relatively straightforward to derive the statistical properties for the resulting parameter estimates.

To illustrate the method in the context of a linear regression problem, where the variation in the response (dependent) variable y can be partly explained by the variation in the independent variables x, consider the model:

$$y = \beta_0 + \beta_1 x + \varepsilon. \tag{2.17}$$

For normality distributed ε, the model can be expressed as $y|x \sim (\beta_0 + \beta_1 x, \sigma^2)$. However, it is not necessary although to assume normality for using the least squares procedure. The term $\beta_0 + \beta_1 x$ is referred as the systematic component of y and ε to as the random component. In practice, if a collection of observations ($x_i = \{x_1, \ldots, x_n\}$) is available one can define the fitted equation as:

$$\hat{y}_i = \beta_0 + \beta_1 x_i \tag{2.18}$$

where β_0 and β_1 are unknown parameters. These two parameters need to be estimated from the data. The least squares estimation provides a way of choosing the values of parameters effectively by minimizing the sum of the squared errors. That is, one chooses the values of β_0 and β_1 that minimize the sum of squared errors:

$$\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2. \tag{2.19}$$

Using some basic mathematical calculus, one can obtain the following estimators for the parameters:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}, \tag{2.20}$$

and

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}, \tag{2.21}$$

where $\bar{x}$ is the mean of the x_i observations and $\bar{y}$ is the mean of the y_i observations.

2.2.2 Maximum Likelihood Estimation

Maximum-likelihood estimation (MLE) is a method of estimating the parameters of a (statistical) model. When applied to a data set and given a statistical model, maximum-likelihood estimation provides estimates for the model's parameters. The method of maximum likelihood corresponds to many well-known estimation methods in statistics. In general, for a fixed set of data and underlying statistical model, the method of maximum likelihood selects the set of values of the model parameters that maximizes the likelihood function. Intuitively, this maximizes the "agreement" of the selected model with the observed data, and for discrete random variables it indeed maximizes the probability of the observed data under the resulting distribution. Maximum-likelihood estimation gives a unified approach to estimation, which is well-defined in the case of the normal distribution and many other problems.

Maximum likelihood is a relatively simple method of constructing an estimator for an unknown parameter vector θ. It was introduced by R. A. Fisher, a great English mathematical statistician, in 1912. Maximum likelihood estimation can be applied in most problems, it has a strong intuitive appeal, and often yields a reasonable estimator of θ. In addition, if the sample data is large, the method will determine a very good estimator of θ. For these reasons, the method of maximum likelihood is probably the most widely used method of estimation in statistics.

Suppose that the random variables $X_1, \ldots, X_n$ form a random sample from a distribution $f(x|\theta)$; if X is continuous random variable, $f(x|\theta)$ is pdf, and if X is discrete random variable, $f(x|\theta)$ is point mass function. We use the given symbol $|$

to represent that the distribution also depends on a parameter θ, where θ could be a real-valued unknown parameter or a vector of parameters. For every observed random sample $x_1, \ldots, x_n$, we define

$$f(x_1, \ldots, x_n|\theta) = f(x_1|\theta) \ldots f(x_n|\theta). \tag{2.22}$$

If $f(x|\theta)$ is pdf, then $f(x_1, \ldots, x_n|\theta)$ is the joint density function; and if $f(x|\theta)$ is pmf, then $f(x_1, \ldots, x_n|\theta)$ is the joint probability. Now we call $f(x_1, \ldots, x_n|\theta)$ as the likelihood function. As we can see, the likelihood function depends on the unknown parameter θ, and it is denoted as $L(\theta)$.

Suppose, for the moment, that the observed random sample $x_1, \ldots, x_n$ came from a discrete distribution. If an estimate of θ must be selected, we would certainly not consider any value of θ for which it would have been impossible to obtain the data $x_1, \ldots, x_n$ that was actually observed. Furthermore, suppose that the probability $f(x_1, \ldots, x_n|\theta)$ of obtaining the actual observed data $x_1, \ldots, x_n$ is very high when θ has a particular value, say $\theta = \theta_0$, and is very small for every other value of θ. Then we would naturally estimate the value of θ to be θ_0. When the sample comes from a continuous distribution, it would again be natural to try to find a value of θ for which the probability density $f(x_1, \ldots, x_n|\theta)$ is large, and to use this value as an estimate of θ. For any given observed data $x_1, \ldots, x_n$, we are led by this reasoning to consider a value of θ for which the likelihood function $L(\theta)$ is a maximum and to use this value as an estimate of θ.

The meaning of maximum likelihood is as follows: we choose the parameter vector that maximizes the likelihood of the observed data. With discrete distributions, the likelihood is the same as the probability. We choose the parameter for the density that maximizes the probability of the data coming from it.

Theoretically, if we had no actual data, maximizing the likelihood function will give us a function of n random variables $X_1, \ldots, X_n$, which we shall call "maximum likelihood estimator" $\hat{\theta}$. When there are actual data, the estimate takes a particular numerical value, which will be the maximum likelihood estimate.

MLE requires us to maximize the likelihood function $L(\theta)$ with respect to the unknown parameter θ. From equation (2.22), $L(\theta)$ is defined as a product of n terms, which is not easy to maximize. Maximizing $L(\theta)$ is equivalent to maximizing $\log L(\theta)$ because log is a monotonic increasing function. We define $\log L(\theta)$ as *log likelihood function*, we denote it as $l(\theta)$, i.e.

$$l(\theta) = \log \prod_{i=1}^{n} f(X_i|\theta) = \sum_{i=1}^{n} \log f(X_i|\theta). \tag{2.23}$$

Maximizing $l(\theta)$ with respect to θ will give us the maximum likelihood estimation. For example, let us consider the method to estimate the parameters μ and σ for the normal density

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\},$$

based on a random sample $X_1, \ldots, X_n$.
First, since we have two unknown parameters, μ and σ, therefore the parameter $\theta = (\mu, \sigma)$ is a vector. We write out the log likelihood function as

$$\begin{aligned} l(\mu, \sigma) &= \sum_{i=1}^{n} \left[-\log \sigma - \frac{1}{2}\log 2\pi - \frac{1}{2\sigma^2}(X_i - \mu)^2\right] \\ &= -n \log \sigma - \frac{n}{2}\log 2\pi - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2. \end{aligned}$$

Setting the partial derivative to be 0, we have

$$\begin{aligned} \frac{\partial l(\mu, \sigma)}{\partial \mu} &= \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \mu) = 0 \\ \frac{\partial l(\mu, \sigma)}{\partial \sigma} &= -\frac{n}{\sigma} + \sigma^{-3}\sum_{i=1}^{n}(X_i - \mu)^2 = 0. \end{aligned}$$

Solving these above equations will give us the MLE for μ and σ:

$$\hat{\mu} = \bar{X} \qquad \text{and} \qquad \hat{\sigma} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

where $\bar{X}$ is the sample mean.

2.3 Value at Risk

One of the main lines of enquiry in this thesis is to see if news enhanced models from the GARCH family can improve risk prediction. We next look at a popular measure of financial risk, namely, Value-at-Risk.

The most traditional measure of risk is volatility. The main problem with volatility, however, is that it does not account for the direction of an investment's movement; a stock can be volatile because it suddenly jumps higher. By assuming investors care about the odds of a really big loss, VaR answers the question, "What is my worst-case scenario?"

In economics and finance, Value-at-Risk (VaR) is the maximum loss not exceeded with a given probability defined as the confidence level, over a given period of time. In other words, a risk management model that calculates the largest possible loss that an institution or other investor could incur on a portfolio. VaR describes the probability of losing more than a given amount, based on a current portfolio. It is commonly used by security houses or investment banks for quantitative risk management of their asset portfolios. Lütkebohmert (2008) points out that VaR does not give any information about the severity of loss by which it is exceeded. Other measures of risk include volatility/standard deviation, semi-variance (or downside risk) and expected shortfall.

VaR is a mathematical model that claims to estimate the maximum future losses expected from a trading portfolio, with a degree of statistical confidence. The VaR of a portfolio with loss variable L at the confidence level q, where $q \in (0,1)$, is given by the smallest number x such that the probability that L exceeds x is not larger than $(1-q)$. The mathematical expression of VaR can be given as:

$$\mathrm{VaR}_q(L) = \inf\{x \in \mathbb{R} : \mathbb{P}(L > x) \leq 1-q\} = \inf\{x \in \mathbb{R} : F_L(x) \geq q\}. \tag{2.24}$$

The distribution function of the loss variable here is $F_L(x) = \mathbb{P}(L \leq x)$. Consequently, VaR is simply a quantile of the loss distribution. In general, VaR can be obtained for different holding periods and confidence levels.

VaR's calculation can be extremely technical, or it can be as simple as looking at a subjective past period and then projecting future risks from there. In particular cases, VaR can very easily produce very low numbers, gravely misrepresenting true exposures. However, underestimating risk is a problem with a particular model, not with VaR. Relying on the past can be false as a quiet past period need not imply future quietness, historical volatility and correlation may mislead you. Also, VaR models can use the "normal probability distribution", which very unrealistically assumes no chance of extreme events. VaR has three parameters:

- The time horizon (period) we are going to analyse (i. e. the length of time over which we plan to hold the assets in the portfolio - the "holding period"). The typical holding period is 1 day, although 10 days are also used. For some problems, even a holding period of 1 year is appropriate.

- The confidence level at which we plan to make the estimate. Popular confidence levels usually are 99% and 95%.
- The unit of the currency which will be used to denominate the VaR.

To give a simple example of VaR, let us consider a portfolio of assets has a one-day VaR of one million GBP with a confidence level of 95%. This means that there is a probability of 0.05 that this portfolio will drop in value by more than one million GBP over a one-day period if there is no trading, i.e. the holdings in the underlying assets in the portfolio do not change over one day period under consideration.

2.3.1 VaR Calculation

A variety of methods exist for estimating VaR. Each model has its own set of assumptions, but the most common assumption is that historical market data is our best estimator for future changes. There are two types of methods: parametric and non-parametric. Common methods include: the historical simulation method, the variance-covariance method and the Monte Carlo simulation.

2.3.1.1 Historical Simulation Method

Historical simulation is the simplest and most transparent method of calculation. This involves running the current portfolio across a set of historical price changes to yield a distribution of changes in portfolio value, and computing a percentile (the VaR). The benefits of this method are its simplicity to implement, and the fact that it does not assume a normal distribution of asset returns. Drawbacks are the requirement for a large market database, and the computationally intensive calculation. It is based on the assumption that history is repeating itself.

Suppose we observe data from day 1 to day t, and r_t is the return of asset on day t, then we get a series of return $\{r_{t+1-i}\}_{i=1}^{N}$. The value at risk with coverage rate (confidence level) p is calculated as the $(100.p)\%$ of the sequence of past asset returns.

$$\mathrm{VaR}_{t+1}^{p} = \mathrm{percentile}\{\{r_{t+1-i}\}_{i=1}^{N}, (100.p)\%\}. \tag{2.25}$$

Unlike other parametric methods, the historical simulation is a non-parametric method; it makes no specific distribution assumption about the distribution of returns. However, the historical simulation implicitly assumes that the distribution of past returns is a good and complete representation of expected future returns. This method also relies on the specified short historical moving window.

2.3.1.2 Variance-Covariance Method

The Variance-Covariance approach is a parametric method. It is based on the assumption that changes in market parameters and portfolio values are normally distributed. As pointed out in Dowd (2006), the assumption of normality is the most basic and straightforward approach and is therefore ideal for simple portfolios consisting of only linear instruments. The advantage of the variance-covariance approach is its simplicity. VaR computation is relatively easy if normality is assumed to prevail, as the standard mathematical properties of the normal distribution can be utilized to calculate VaR levels. In addition, normality allows easy translatability between different confidence levels and holding periods. The Variance-Covariance method of one step ahead VaR forecast is

$$VaR_{(t|t-1)} = -\alpha \times \sigma_t, \tag{2.26}$$

where α is 1.645 for 95% confidence level and 2.33 for 99% confidence level and σ_t is the estimated volatility forecasted at time $t-1$ by the candidate model. Here, the mean return is assumed to be zero for short forecasting horizons. Further details can be found in Linsmeier et al. (1996).

Despite the ease of implementation of this method, the assumption of normality also causes problems. Jorion (1997) points out that most financial assets are known to have "fat tailed" return distributions, meaning that in reality extreme outcomes are more probable than normal distribution would suggest. As a result, VaR estimates may be understated.

Problems grow even bigger when the portfolio includes instruments, such as options, whose returns are non-linear functions of risk variables. One solution to this issue is to take first order approximation to the returns of these instruments and then use the linear approximation to compute VaR. This method is called *delta-normal* approach. However, Dowd (2006) showed that the shortcoming of *delta-normal* method is that it only works if there is limited non-linearity in the portfolio.

2.3.1.3 Monte Carlo Simulation Method

Monte Carlo simulation is another non-parametric method. It is the most popular approach when there is a need for a sophisticated and powerful VaR system, but it is also by far the most challenging one to implement. As explained in Jorion (1997), the Monte Carlo simulation process can be described in two steps. First, stochastic processes for financial variables are specified and correlations and volatilities are estimated on the basis of market or historical data. Second, a large number of sample price paths for all financial variables are simulated using an appropriate model for price dynamics (e.g. Geometric Brownian motion). These price realizations are then compiled to a joint distribution of returns, from which VaR estimates can be calculated. The strength of Monte Carlo simulation is that no assumptions about normality of returns have to be made. Even though parameters are estimated from historical data, one can easily bring subjective judgements and other information to improve forecasted simulation distributions. Damodaran (2007) outlined that the method is also capable of covering non-linear instruments, such as options. In addition to these advantages, Jorion reminds that Monte Carlo simulation generates the entire distribution and therefore it can be used, for instance, to calculate losses in excess of VaR.

The most significant problem with Monte Carlo approach, as mentioned in Jorion (1997), is its computational time. The method requires a lot of resources, especially with large portfolios. As a consequence, the implementation may turn out to be expensive. A potential weakness is also model risk, which arises due to wrong assumptions about the pricing models and underlying stochastic processes. If these are not specified properly, VaR estimates will be distorted. Moreover, Dowd (1998) points out that complicated procedures associated with this method require special expertise.

2.3.2 Assessment of VaR Methods

The previous section discussed the common shortcomings of the different VaR methods. Let us now turn the focus towards the general criticism that has been raised against VaR as a risk management tool.

The concept of VaR is very simple but this is also one of the main sources of critique. VaR reduces all the information down to a single number, meaning the loss of potentially important information. For instance, VaR gives no information on the extent of the losses that might occur beyond the VaR estimate. As a result, VaR estimates may lead to incorrect interpretations of prevailing risks. Tsai (2004) points out that one thing particularly important to realize is that portfolios with the same VaR do not necessarily carry the same risk. Longin (2001) suggests a method called Conditional VaR to deal with this problem. Conditional VaR measures the expected value of the loss in those cases where VaR estimate has been exceeded.

VaR has also been criticized for its narrow focus. Damodaran (2007) shows that VaR in its conventional form is unable to account for any other risks than market risk. However, VaR has been extended to cover other types of risks. For instance, Jorion (1988) points out that Monte Carlo simulation can handle credit risks to some extent. VaR has also problems in estimating risk figures accurately for longer time horizons as the results quickly deteriorate when moving e.g. from monthly to annual measures. Further criticism has been presented by Kritzman and Rich (2002), who point out that VaR considers only the loss at the end of the estimation period, but at the same time many investors look at risk very differently. They are exposed to losses also during the holding period but this risk is not captured by normal VaR models. To take into account for this, the authors suggest a method called continuous Value at Risk.

Many economists argue that history is not a good predictor of the future events. For example, Damodaran (2007) points out that all VaR methods still rely on historical data, at least to some extent. In addition, every VaR model is based on some kinds of assumptions which are not necessarily valid in any circumstances. Tsai (2004) emphasizes that VaR estimates should therefore always be accompanied by other risk management techniques, such as stress testing, sensitivity analysis and scenario analysis in order to obtain a wider view of surrounding risks.

2.3.3 Model Backtesting

Financial firms are often obliged to use VaR (see, e.g. Basel II in Supervision (2011)). Firms that use VaR as a risk disclosure or risk management tool are facing growing pressure from internal and external parties such as senior management, regulators, auditors, investors, creditors, and credit rating agencies to provide estimates of the accuracy of the risk models being used.

Users of VaR realized early that they must carry out a cost-benefit analysis with respect to the VaR implementation. A wide range of simplifying assumptions is usually used in VaR models (distributions of returns, historical data window defining the range of possible outcomes, etc.), and as the number of assumptions grows, the accuracy of the VaR estimates tends to decrease.

As the use of VaR extends from pure risk measurement to risk control in areas such as VaR-based Stress Testing and capital allocation, it is essential that the VaR provide accurate information, and that someone in the organization is accountable for producing the best possible risk estimates. In order to ensure the accuracy of the forecasted VaR, risk managers should regularly backtest the risk models being used, and evaluate alternative models if the results are not entirely satisfactory.

Backtesting is a statistical procedure where actual profits and losses are systematically compared to corresponding VaR estimates. For example, if the confidence level used for calculating daily VaR is 99%, we expect an exception to occur once in every 100 days on average. In the backtesting process we could statistically examine whether the frequency of exceptions over some specified time interval is in line with the selected confidence level. These types of tests are known as tests of *unconditional coverage*. They are straightforward tests to implement since they do not take into account for when the exceptions occur.

In theory, however, a good VaR model not only produces the "correct" amount of exceptions but also exceptions that are evenly spread over time (i.e. are independent of each other). Clustering of exceptions indicates that the model does not accurately capture the changes in market volatility and correlations. Tests of *conditional coverage* therefore examine also conditioning, or time variation, in the data. Further information about tests of conditional and unconditional coverage can be found in Jorion (2001).

This section aims to provide an insight into different methods for backtesting a VaR model. Keeping in mind that the aim of this research is in the empirical study, the focus is on those backtests that will be applied later in the empirical part. The tests include Kupiec's proportion of failures-test proposed in Kupiec (1995) and Christoffersen's interval forecast test proposed in Christoffersen (1998).

2.3.3.1 Unconditional Coverage

The most common test of a VaR model is to count the number of VaR exceptions, i.e. days (or holding periods of other length) when portfolio losses exceed VaR estimates. If the number of exceptions is less than the selected confidence level would indicate, the system overestimates risk. On the contrary, too many exceptions signal underestimation of risk. Naturally, it is rarely the case that we observe the exact amount of exceptions suggested by the confidence level. It therefore comes down to statistical analysis to study whether the number of exceptions is reasonable or not, i.e. will the model be rejected or not.

Denoting the number of exceptions as x and the total number of observations as T, we may define the *failure rate* as x/T. In an ideal situation, this rate would reflect the selected confidence level. For instance, if a confidence level of 99% is used, we have a null hypothesis that the frequency of tail losses is equal to $p = (1-c) = 1-0.99 = 1\%$. Assuming that the model is accurate, the observed failure rate x/T should act as an unbiased measure of p, and thus converge to 1% as sample size is increased as outlined in Jorion (2001).

Each portfolio observation/return either produces a VaR exception or not. This sequence of "successes and failures" is commonly known as Bernoulli trial. The number of exceptions x follows a Binomial probability distribution:

$$f(x) = \binom{T}{x} p^x (1-p)^{T-x}. \tag{2.27}$$

As the number of observations increases, the Binomial distribution can be approximated with a normal distribution:

$$z = \frac{x - p^T}{\sqrt{p(1-p)^T}} \approx N(0,1), \tag{2.28}$$

where p^T is the expected number of exceptions and $p(1-p)^T$ the variance of exceptions. By utilizing this Binomial distribution we can examine the accuracy of the VaR model, as discussed in the next paragraph.

Kupiec Test:

The most widely known test based on failure rates has been suggested by Kupiec (1995). Kupiec's test, also known as the POF-test (proportion of failures), measures whether the number of exceptions is consistent with the confidence level. Under null hypothesis of the model being "correct", the number of exceptions follows the binomial distribution as discussed earlier. Let x be the number of times the portfolio loss is worse than the true Value-at-Risk in a sample of size T. Then the number of VaR exceptions has a Binomial distribution, $x \sim B(T;p)$. This test has a null hypothesis that the failure rate of the VaR is equal to the chosen percentage of losses (e.g. 10%, 5% or 1%), such that

$$H_0 : p = 1 - c,$$

where p is the probability of the failure/exceedance and c is the confidence level of VaR. The POF-test is best conducted as a likelihood-ratio (LR) test. Hence, the only information required to implement a POF-test is the number of observations (T), number of exceptions (x) and the confidence level (c). The test statistic is given by:

$$LR_{POF} = -2\log\left[(1-c)^x (c)^{T-x}\right] + 2\log\left[\left(1-\left[\frac{x}{T}\right]\right)^{T-x}\left(\frac{x}{T}\right)^x\right]. \tag{2.29}$$

The likelihood ratio (LR_{POF}) is asymptotically χ^2 (Chi-squared) distributed with one degree of freedom under the null hypothesis that p is the true probability the VaR is exceeded. If the value of the LR_{POF} -statistic exceeds the critical value of the χ^2 distribution, the null hypothesis will be rejected and the model is deemed as inaccurate. According to Dowd (2006), the confidence level (i.e. the critical value) for any test should be selected to balance between type I and type II errors. Type I error refers to the possibility of rejecting a correct model and type II error to the possibility of not rejecting an incorrect model. It is common to choose some arbitrary confidence level, such as 95%, and apply this level in all tests. A level of this magnitude implies that the model will be rejected only if the evidence against it is fairly strong.

Probability Level p	VaR Confidence Level	Non-rejection Region for Number of Failures N		
		T=255 days	T=510 days	T=1000 days
0.01	99%	N < 7	1 < N < 11	4 < N < 17
0.025	97.5%	2<N<12	6<N<21	15<N<36
0.05	95%	6<N<21	16<N<36	37<N<65
0.075	92.5%	11<N<28	27<N<51	59<N<92
0.1	90%	16<N<36	38<N<65	81<N<120

Table 2.1: Non-rejection regions for Kupiec's test under different confidence levels

Table 2.1 displays Non-rejection regions for the POF-test. The figures show how the power of the test increases as the sample size gets larger. Clearly, the smaller the left tail probability, the more difficult it gets to confirm deviations, especially when the evaluation sample size is small.

Thus, as outlined in Jorion (2001), with more data we are able to reject an incorrect model more easily. Kupiec's POF-test is hampered by two shortcomings. First, the test is statistically weak with sample sizes consistent with current regulatory framework (one year). This lack of power has already been recognized by Kupiec himself [1]. Secondly, the POF-test considers only the frequency of losses and not the time when they occur. As a result, it may fail to reject a model that produces clustered exceptions. Thus, Campbell (2005) confirmed that model backtesting should not rely solely on tests of unconditional coverage.

2.3.3.2 Conditional Coverage

The unconditional coverage tests, such as the POF-test, focus only on the number of exceptions. In theory, however, we would expect these exceptions to be evenly spread over time. Therefore, as discussed in Finger (2006), good VaR models are capable of reacting to changing volatility and correlations in a way that exceptions occur independently of each other, whereas bad models tend to produce a sequence of consecutive exceptions.

Clustering of exceptions is something that VaR users want to be able to detect since large losses occurring in rapid succession are more likely to lead to disastrous events than individual exceptions taking place every now and then as discussed in Christoffersen and Pelletier (2004). Tests of conditional coverage deal with this problem by not only examining the frequency of VaR violations but also the time when they occur. In the following paragraph, the Christoffersen test will be presented.

Christoffersen's Interval Forecast Test:

Probably the most widely known test of conditional coverage has been proposed by Christoffersen (1998). He uses the same log-likelihood testing framework as Kupiec, but extends the test to include also a separate statistic for independence of exceptions. In addition to the correct rate of coverage, his test examines whether the probability of an exception on any day depends on the outcome of the previous day. The testing

[1] The power of a hypothesis test is the probability of rejecting the null hypothesis H_0 when the alternative hypothesis H_1 is true. Availability of only a small number of samples leads to a loss of powerful; see Weiss and Hassett (1999) for more details on power analysis.

procedure described below is explained, for example, in Jorion (2001), Campbell (2005), Dowd (2006) and in greater detail in Christoffersen (1998).

The test is carried out by first defining an indicator variable that gets a value of 1 if VaR is exceeded and value of 0 if VaR is not exceeded:

$$I_t = \begin{cases} 1 & \text{if violation occurs} \\ 0 & \text{if no violation occurs.} \end{cases}$$

Then define n_{ij} as the number of days when condition j occurred assuming that condition i occurred on the previous day. To illustrate, the outcome can be displayed in a 2×2 contingency table:

	$I_{t-1} = 0$	$I_{t-1} = 1$	
$I_t = 0$	n_{00}	n_{10}	$n_{00} + n_{10}$
$I_t = 1$	n_{01}	n_{11}	$n_{01} + n_{11}$
	$n_{00} + n_{01}$	$n_{10} + n_{11}$	N

In addition, let π_i represent the probability of observing an exception conditional on state i on the previous day:

$$\pi_0 = \frac{n_{01}}{n_{00}+n_{01}}, \qquad \pi_1 = \frac{n_{11}}{n_{10}+n_{11}}, \qquad \pi = \frac{n_{01}+n_{11}}{n_{00}+n_{01}+n_{10}+n_{11}}.$$

If the model is accurate, then an exception today should not depend on whether or not an exception occurred on the previous day. In other words, under the null hypothesis the probabilities π_0 and π_1 should be equal. The relevant test statistic for independence of exceptions is a likelihood-ratio:

$$LR_{ind} = -2\ln\left(\frac{(1-\pi)^{n_{00}+n_{10}}\pi^{n_{01}+n_{11}}}{(1-\pi_0)^{n_{00}}\pi_0^{n_{01}}(1-\pi_1)^{n_{10}}\pi_1^{n_{11}}}\right). \tag{2.30}$$

By combining this independence statistic with Kupiec's POF-test we obtain a joint test that examines both properties of a good VaR model, the correct failure rate and independence of exceptions, i.e. conditional coverage:

$$LR_{cc} = LR_{POF} + LR_{ind}.$$

LR_{cc} is also χ^2 (chi-squared) distributed, but in this case with two degrees of freedom since there are two separate LR-statistics in the test. If the value of the LR_{cc} -statistic is lower than the critical value of the χ^2 distribution, the model passes the test. Higher values lead to rejection of the model.

Christoffersen's framework allows examining whether the reason for not passing the test is caused by inaccurate coverage, clustered exceptions or even both. This evaluation can be done simply by calculating each statistic, LR_{POF} and LR_{ind}, separately and using the χ^2 distribution with one degree of freedom as the critical value for both statistics. Campbell (2005) reminds that in some cases it is possible that the model passes the joint test while still failing either the independence test or the coverage test. Therefore it is advisable to run the separate tests even when the joint test yields a positive result.

In the subsequent chapter, we will investigate the impact of quantitative news sentiment on the volatility of stock market returns of two different markets using four predictive models. The estimated models will be compared in terms of their goodness of fit to data as well as their out of sample predictive ability, each model under two distributional assumptions: Gaussian and student-t distributions. The aim is to find out whether the proposed model, which will be introduced in the next chapter, has a superior forecasting ability compared to the existing GARCH-type models. Furthermore, the study will carry out an empirical application to the VaR with the different volatility estimates obtained by the models.

Chapter 3

Enhancing the Prediction of Volatility Using News Sentiment

3.1 Introduction and Background

Financial time series plays a fundamental role in modelling and forecasting stock return volatility in the financial markets. Accurate prediction of asset volatility is important for market participants, such as option traders and portfolio managers, for improved risk management and asset allocation. Volatility prediction is also relevant for policy makers since excess volatility reflects increased uncertainty and may adversely affect growth prospects. Hence ways to improve volatility forecasting using news sentiment are worth investigating.

Over the last few decades, the investigation of modelling and analysing the temporal behaviour in the conditional variance of financial time series has been considered by many researchers. A large number of academics have focused on the analysis and forecasting of stock return volatility. As a result, a number of models have been proposed to estimate and forecast the conditional volatility of financial instruments, of which the most popular are the conditional heteroscedastic models. The most well-known approaches are GARCH type models, for example, as described earlier in chapter 2, GARCH of Bollerslev (1986), EGARCH of Nelson (1991), GJR-GARCH of Glosten et al. (1993), TGARCH of Zakoian (1994) and GRS-GARCH of Gray (1996) amongst others, which cover symmetric and asymmetric effects of news in volatility. On the other hand, the impact of news on financial markets has influenced many researchers to explore the relationship between published news and financial instruments behaviours. For instance, Mitchell and Mulherin (1994) observed that the relation between news and market activity is not particularly strong and the patterns in news announcements do not explain the day-of-the-week seasonalities in market activity. Andersen (1996) showed that different types of news have a different effect on the conditional stock volatility. Crouhy and Rockinger (1997) confirmed that volatility rises more in response to bad news than to good news. Li and Engle (1998) studied the effect of the macroeconomic public announcements and revealed the heterogeneous persistence from scheduled news versus non-scheduled news, emphasising how the prediction of volatility can be improved by considering only those news corresponding to scheduled events. Kalev et al. (2004) gave evidence that news arrivals display a very strong pattern of autocorrelation. Cousin and de Launois (2006) considered the daily number of press releases on a stock (news intensity) as the most appropriate explanatory variable that can improve GARCH model. Tetlock (2007) explored the interactions between investor sentiment and stock market. Tetlock (2010) found four patterns in post-news returns and trading volume that are consistent with the asymmetric information model's predictions. The empirical results of Riordan et al. (2013) confirmed that negative news messages induce stronger market reactions than positive ones.

Despite the fact that the early generation of GARCH models provide good volatility forecasts and simple parametrization (see, for example, Hien and Thanh (2008), Alberg et al. (2008) and Liu et al. (2009) among others), the models cannot fully capture the asymmetric behaviour between stock returns and volatility such as volatility clustering and leverage effect that was discovered by Black (1976). Moreover, as discussed in Andersen and Bollerslev (1997), GARCH models give good in-sample results in terms of fit to the data, but often give bad volatility forecasts out of sample. Therefore, it is of interest to further investigate and utilize exogenous sources of information beside the market asset price data to improve the predictive power of the existing GARCH type models. Exogenous information, in addition to the traditional conditional heteroscedastic model structure, can help us to better capture the market behaviour and therefore improving the prediction of the financial market's reactions. Examples of improvements in the forecasting ability of volatilty models when using trading volume as an exogenous input have been reported in Sharma et al. (1996), Aragó and Nieto (2005), Xiao et al. (2009), Ashok and Rahul (2011) and Sidorov et al. (2013). Engle et al. (2001) have considered interest rate levels as an exogenous input. All the aforementioned studies show improvement of the model forecasting ability by including an appropriate exogenous variable in the model structure.

Financial markets incorporate new information from a variety of news sources (such as newscasts, articles or announcements) rapidly and this information can influence the asset price movements. The strong relationship between news flows and stock prices fluctuations in the financial markets, as well as the birth of news analytics providers such as RavenPack (2014) and Thomson Reuters (2010), who convert the textual input of news stories into quantitative sentiment scores, have encouraged numerous researches to investigate news sentiment analysis for systematic trading. News analytics data has come out in the last few years as a valuable new data source for investors to create systematic trading models. There are a limited number of studies that exploit news sentiment scores to enhance stock volatility predictions. Over the last few years, much of the attention has been focused on equities. For example, Mitra et al. (2008) demonstrated that news has a significant impact on the asset's volatility and its inclusion in the GARCH model improves the forecast. Hafez (2009) looked at how company-specific news impacts on equities. Hafez (2013) also has focused on trading equity indices through the construction of sentiment indices. More recent, Yu (2014) and Yu et al. (2015) enhanced the GARCH model by adding news impact scores as external regressors in the model equation. They split the news impact according to news sentiment signs, *i.e.* depending on whether the impact was positive or negative.

The study reported in this chapter utilizes daily aggregated news sentiment scores using the method proposed in Yu (2014) to enhance the predictive ability of symmetric GARCH model. Here, the RavenPack news sentiment score is utilized as a quantitative proxy for news sentiment. This proxy is then used as an exogenous term in a modified version of the GARCH model proposed by Bollerslev (1986). The proposed modified version of the GARCH model is called *first* news augmented GARCH model (NA1-GARCH). The new model incorporates the impact of news into the volatility prediction model in a way which is meaningful from an economic point of view. The objective in this study is not to investigate whether a GARCH model provides the most accurate and robust forecasts, but it is to see whether modified versions of the GARCH model can consistently outperform predictions from traditional and more parsimonious models.

In this study, four volatility models have been calibrated on datasets from two different financial markets via maximum likelihood estimation. The models are simple GARCH, threshold GARCH (TGARCH), exponential GARCH (EGARCH) and the proposed NA1-GARCH model. The estimated models are then compared in terms of their in-sample fit to data and out-of-sample predictive ability, where the forecasts of the different models are compared to realized volatility which was estimated using daily stock returns. Many other variants of GARCH models exist; e.g., see AGARCH (Asymmetric GARCH) of Engel (1990), CGARCH (Component GARCH) of Lee and Engle (1993), GJR-GARCH of Glosten et al. (1993), FIGARCH (Fractionally

Integrated GARCH) of Baillie et al. (1996), PGARCH (Periodic GARCH) of Bollerslev and Ghysels (1996), ATGARCH (Asymmetric Threshold GARCH) of Crouhy and Rockinger (1997), among others. This study has restricted the discussion to these four models: a 'benchmark' GARCH, two asymmetric variants of GARCH (TGARCH and EGARCH) and first news augmented GARCH (NA1-GARCH). Another reason for not extending the study into other GARCH models is the existing evidence in the literature (see, e.g. Hansen and Lunde (2005)) that it is hard to beat GARCH(1,1) model in terms of its forecasting ability.

The remainder of the chapter is organized as follows. Section 2 provides descriptions of two streams of time series that have been used for numerical experiments, the market data and news meta data of two indices; FTSE100 and S&P500. Section 3 presents the proposed volatility model (NA1-GARCH) that enables one to use the news sentiment score to improve the predictive ability of GARCH model, as well as explains the calibration of the used models by using the maximum likelihood estimation, each under two different distributional assumptions: Gaussian distributed and Student's t-distribution. Section 4 explains the methodology employed in comparing the performance of the models and presents the results of the empirical investigation of the estimated models for the datasets. Finally, Section 5 concludes the chapter.

3.2 Data

In this study, two streams of time series data have been used: (i) Market data, which is given on a daily basis as asset closing prices, and (ii) News meta data as supplied by RavenPack.

3.2.1 Market Data

The time series of market data used for modelling volatility in this study is the stock market daily closing prices of two indices: FTSE100 and S&P500, from 3 January 2005 to 23 November 2015 for a total of 3976 data points (observations) for each index, which was obtained from Datastream. This data set was further split into 24 overlapping data subsets, each with 750 consecutive data points. Each of the data subset starts on the first trading date of three different months each year from 2005 to 2012. For example, in 2005, the first trading date of January, April and August are chosen to be the start of three different data subsets (each with 750 data points);

please see table 3.1 for more information about how the data was split into multiple datasets. Out of 750 points in each subset, the first 500 points are used for parameter estimation and the subsequent 250 points are used for out-of-sample comparison of the models using comparison measures described later in section 3.5.2. The choice of the datasets in this way is due to the purpose to test the performance of the models for different economic periods (including a recession in 2008) and for different markets (UK and USA).

FTSE100				SP500			
Datasets	In Sample	Out Of Sample		Datasets	In Sample	Out Of Sample	
	Start date	Start date	End date		Start date	Start date	End date
FTSE1	2005-01-03	2006-12-04	2007-11-16	**SP1**	2005-01-03	2006-12-27	2007-12-24
FTSE2	2005-04-01	2007-03-02	2008-02-14	**SP2**	2005-04-01	2007-03-28	2008-03-25
FTSE3	2005-08-01	2007-07-02	2008-06-13	**SP3**	2005-08-01	2007-07-27	2008-07-23
FTSE4	2006-02-01	2008-01-02	2008-12-16	**SP4**	2006-02-01	2008-01-29	2009-01-23
FTSE5	2006-04-03	2008-03-03	2009-02-13	**SP5**	2006-04-03	2008-03-31	2009-03-25
FTSE6	2006-09-01	2008-08-01	2009-07-16	**SP6**	2006-09-01	2008-08-28	2009-08-25
FTSE7	2007-05-01	2009-03-31	2010-03-15	**SP7**	2007-05-01	2009-04-24	2010-04-21
FTSE8	2007-07-02	2009-06-01	2010-05-14	**SP8**	2007-07-02	2009-06-25	2010-06-22
FTSE9	2007-12-03	2009-11-02	2010-10-15	**SP9**	2007-12-03	2009-11-25	2010-11-22
FTSE10	2008-03-03	2010-02-01	2011-01-14	**SP10**	2008-03-03	2010-02-25	2011-02-18
FTSE11	2008-08-01	2010-07-02	2011-06-16	**SP11**	2008-08-01	2010-07-28	2011-07-22
FTSE12	2008-10-01	2010-09-01	2011-08-16	**SP12**	2008-10-01	2010-09-27	2011-09-21
FTSE13	2009-01-05	2010-12-06	2011-11-18	**SP13**	2009-01-05	2010-12-29	2011-12-22
FTSE14	2009-06-01	2011-05-02	2012-04-13	**SP14**	2009-06-01	2011-05-24	2012-05-18
FTSE15	2009-09-01	2011-08-02	2012-07-16	**SP15**	2009-09-01	2011-08-25	2012-08-21
FTSE16	2010-02-01	2012-01-02	2012-12-14	**SP16**	2010-02-01	2012-01-25	2013-01-23
FTSE17	2010-05-03	2012-04-02	2013-03-15	**SP17**	2010-05-03	2012-04-25	2013-04-24
FTSE18	2010-10-01	2012-08-31	2013-08-15	**SP18**	2010-10-01	2012-09-25	2013-09-24
FTSE19	2011-03-01	2013-01-29	2014-01-13	**SP19**	2011-03-01	2013-02-26	2014-02-21
FTSE20	2011-06-01	2013-05-01	2014-04-15	**SP20**	2011-06-01	2013-05-29	2014-05-23
FTSE21	2011-11-01	2013-10-01	2014-09-22	**SP21**	2011-11-01	2013-10-29	2014-10-24
FTSE22	2012-03-01	2014-01-30	2015-01-26	**SP22**	2012-03-01	2014-02-27	2015-02-24
FTSE23	2012-07-02	2014-06-06	2015-06-02	**SP23**	2012-07-02	2014-06-30	2015-06-25
FTSE24	2012-12-03	2014-11-10	2015-11-04	**SP24**	2012-12-03	2014-11-26	2015-11-23

Table 3.1: In-Sample and Out-Of-Sample dates for all datasets

In this study, daily returns (r_t) were calculated as the continuously compounded returns which are the first difference in logarithm of closing prices of the index of successive days:

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right) \tag{3.1}$$

where P_t and P_{t-1} are the daily closing price of the index at current and previous days, respectively.

3.2.2 News Data and its Impact Measurement

News analytics have dramatic impact on the financial markets in recent years. Most of the quantitative firms are utilizing some kind of machine readable news-feed data to build up or enhance their strategies. News analytics have also been of interest for a large number of academics. However, news data is unstructured data in textual form. It is hard to interpret whether a specific news item has positive or negative impact on the market, if it is relevant to more than a particular company, and whether it is new information or a repetition of a previous story. News analytic vendors, such as Ravenpack and Thomson Reuters, convert the unstructured news data into structured data using various proprietary algorithms. These algorithms use data mining to identify information such as: the company (or companies) referred to in each news item, the relevance of the item to the company, whether the news item is positive or negative for the company etc. They assign for each news item a time-stamp, at which the news item was released, and a unique identifier for each company that was mentioned in the news item. Then characteristics such as relevance, event novelty score (ENS), event sentiment score (ESS) are given to each identified news item. These characteristics differ in values for each mentioned company or financial instrument.

In this study, RavenPack news analytics metadata is used. News stories in a RavenPack database are classified into five topics, namely: business, economy, environment, politics and society. Each record (row) in the news analytics database has 46 fields including a time-stamp, reference identifier, relevance, novelty, sentiment, and a unique identifier for the analysed news story. The time-stamp is represented to the second using the Coordinated Universal Time (UTC). Relevance takes a score between zero to 100, which represents how strongly related the company is to the analysed news story. A relevance score of 100 signals that the company plays a key role in the news story and the content is most relevant for this company. Novelty takes a score between one to 100, which indicates how new the analysed news story is within a 24-hour time window. A novelty score of 100 refers to the first story reporting a categorized event and is considered to be the most novel. Sentiment takes a score between zero and 100 that represents the news sentiment for a given company. RavenPack produces an individual ESS for each news item which is a relative number that describes the degree of positivity and negativity in a piece of news, where 0 is the most negative score and 100 is the most positive score. The full explanation of the RavenPack data fields can be found in RavenPack (2014) news analytics manual.

The trading day for the London Stock Exchange, where FTSE100 index components are listed, starts at 08:00 hours and ends at 16:30 hours thus the total number of minutes in a trading day is 510. The trading day for NASDAQ and New York Stock Exchanges, where S&P500 index components are listed, starts at 09:30 hours and ends at 16:00 hours thus the total number of minutes in a trading day is 390. Therefore, any news arriving overnight (after 16:30 UTC for FTSE100 index components and 16:00 EST for S&P500 index components) or during the weekend or holidays is bucketed into the next business day first minute. Hence, the assumption is taken that the impact of such overnight and weekend news is incorporated into prices the following trading day.

Specific filters on RavenPack fields have been used to construct the market sentiment proxies that impact FTSE100 and S&P500 indices. In particular, the event sentiment scores have been filtered using fields such as relevance and novelty. First, the news metadata is filtered for the chosen indices (FTSE100 and S&P500), the news item were selected under the filter of relevance score of 100, and novelty scores of 70 or more in the dataset. Then, the event sentiment scores, which range from 0 to 100, have been transformed into a scaled sentiment score in the range from -50 to +50, as it is found that such a derived single score provides a relatively better interpretation of the mood of the news item.

For this study, low frequency market data (closing daily prices) has been used. Thus, the news impact score was taken at the last minute in the trading day, which was the aggregation of all news data from 08:00 to 16:30 for FTSE100 index (and 09:30 to 16:00 for S&P500 index). This data actually is normalized or scaled in the range between 0 and 100, and represented by two data streams, positive and negative impact scores.

In order to drive suitable news impact terms (positive and negative impact scores) and utilize them as proxies of good and bad news, the following points have to be taken in consideration:

(i) An expression has to be found which describes the attenuation of the news sentiment score.

(ii) Bad news is more affective on the financial markets than good news.

(iii) The impact of a news item does effect markets at the time of release and persists over a finite period of time.

(iv) The impact of a news item will decay over the time.

To account for these points, this study has used the decay model technique that was proposed by Yu (2014) which reflect the instantaneous impact of news releases and the decay of this impact over a subsequent period of time. The technique combines exponential decay and accumulation of the sentiment score over a given time bucket under observation. The technique can be summarised as follows:

- Let $\mathbf{N}^i_\tau$ be a filtered news item based on accepted relevance and novelty scores, where i is the asset (asset class, index, ...) and τ is the time stamp, where $\tau = 1, \cdots T$ and T is the last time bucket in the day.
- As mentioned earlier, a news sentiment score between 0 and 100 is provided by Ravenpack for each time-stamped news. We transfer this score into -50 and +50 and denote it by $\mathcal{S}_\tau(\mathbf{N}^i_\tau)$, such that $\mathcal{S}_\tau(\mathbf{N}^i_\tau) \in [-50, 50]$ for each news item $\mathbf{N}^i_\tau$.
- The impact of a particular news item exponentially decays over time. We model the impact score $\mathbf{I}_t(\mathbf{N}^i_\tau)$ at time $t \geq \tau$ as:

$$\mathbf{I}_t(\mathbf{N}^i_\tau) = \underbrace{\mathcal{S}_\tau(\mathbf{N}^i_\tau)}_{\text{Sentiment score}} \underbrace{e^{-\lambda(t-\tau)}}_{\text{Decay}}, \quad t \geq \tau \tag{3.2}$$

where t represents the closing time of stock market (e.g. 16:30 for London Stock Exchange) and τ is the time when news breaks. Therefore $(t - \tau)$ measures the difference between the time when a news happened and the closing time of the stock market. λ is the exponent which determines the decay rate. This exponential decay effect causes a news story to have only half of the initial impact left after a time span of 90 minutes. In this study, to determine the value of λ three different decay rates have been considered; 0.5, $\frac{2}{3}$ and 0.75. For each decay rate, six values of the decay duration (15, 30, 45, 60, 90, 120 minutes) have been tested. The movement of cumulated sentiment scores, which incorporates the two variables, is compared to the movement of bid price. After testing a combination of experiments, it is found that fixing the decay rate to 0.5 and decay duration to 90 minutes are most suitable values from the values which were tested. The justification of such conclusion is that the movement of cumulated sentiment scores is most synchronised with the bid price movement. A similar effect was also observed in Yu (2014), Arbex-Valle et al. (2013) and Yu et al. (2015). Further details about the comparison method of the variables with bid price can be found in Yu (2014). Therefore, we specify the decay rate to 0.5 and decay duration to 90 minutes to calculate the value of λ. Hence, λ can be determined from the expression: $1 * e^{-\lambda(90)} = \frac{1}{2} \implies \lambda = 0.007701635$.

- Instead of simply aggregating impact of news items with positive and negative impact, they are kept separated so that positive and negative effects shall not cancel out. Cancellation reduces the news flow and can lead to misinterpretation.

- We define $\mathcal{I}_\mathcal{P}$ and $\mathcal{I}_\mathcal{N}$ to be sets of news items with positive and negative impact score, respectively. In particular,

$$\mathcal{I}_\mathcal{P} = \left\{ \mathbf{N}_\tau^i \mid \mathbf{I}_t(\mathbf{N}_\tau^i) \geq \theta \right\} \quad \text{and} \quad \mathcal{I}_\mathcal{N} = \left\{ \mathbf{N}_\tau^i \mid \mathbf{I}_t(\mathbf{N}_\tau^i) \leq -\theta \right\} ,$$

  for a threshold $\theta > 0$.

- In order to obtain daily positive and negative impact scores, for each day all the positive and negative impact scores are aggregated separately for asset i at time point T, where T is the last minute of the trading day. The mathematical expression is as follows:

$$\mathcal{P}_T^i = \sum_{\mathbf{N}_\tau^i \in \mathcal{I}_\mathcal{P}} \mathbf{I}_t(\mathbf{N}_\tau^i) \quad \text{and} \quad \mathcal{N}_T^i = \sum_{\mathbf{N}_\tau^i \in \mathcal{I}_\mathcal{N}} \mathbf{I}_t(\mathbf{N}_\tau^i).$$

- Finally, the time series of positive and negative impact scores have been transformed into a scaled scores in the range 0 to 100 for positive scores and -100 to 0 for negative scores.

The separation of the positive and negative sentiment scores is only logical as this avoids cancellation effects and the resultant misinterpretation of news. In addition, note that although news data is intra-day, it is not conformed to a regular time scale thus there is not a guaranteed number of data points within each day for a particular index. More details can be found in Yu et al. (2015).

3.3 First News Augmented GARCH Model (NA1-GARCH)

In many cases, the basic GARCH conditional variance equation (2.5) under normality assumption provides a reasonably good model for analysing financial time series and estimating conditional volatility. However, in some cases there are aspects of the model which can be improved so that it can better capture the characteristics and dynamics of a particular time series.

Before introducing a new model structure to improve the volatility prediction of GARCH(1,1) model using news data, it is worth recalling that trading on financial markets is strongly influenced by company-specific, macroeconomic or political information flows. As a result, markets react sensitively to news, which is announced on a regular and irregular basis, and news events appear to affect stock return volatility quickly, suggesting that the market incorporates information quickly. In particular, volatility tends to be higher in a falling market than in a rising market. Commercial news analytics providers such as RavenPack and Thomson Reuters have started to automatically track and monitor relevant information on tens of thousand of companies and quantifying the content of news articles about them, and allowing for the measurement of the reaction to positive, neutral and negative news.

The objective in this study is to develop a volatility prediction model which enables us to use the information content of news events in order to improve the volatility prediction using GARCH (1,1) model. The first step is to construct news impact scores as described in subsection 3.2.2, to be used as proxies of news events in the new model structure. To accomplish this, RavenPack's news analytics database is used and some of its quantitative sentiment scores are exploited. Since the financial markets are mainly sensitive to good and bad news, only positive and negative news sentiments are taken into consideration and it is assumed that neutral news does not have any effect on stock return volatility.

The model structure needs to reflect the following economic realities: positive and negative news impacts the volatility differently, positive news tends to reduce volatility whereas negative news tends to increase volatility, and finally the impact of news on return's volatility of an asset decays relatively slowly (more as a power law than as an exponential).

Let $\mathcal{P}_t$ and $\mathcal{N}_t$ represent positive and negative news impact scores at time t

respectively. Keeping in mind the economic realities mentioned above, the following function is defined as a scaling factor to GARCH(1,1) model:

$$f(\mathcal{P}_t, \mathcal{N}_t) = \left[\frac{1}{\lambda + \left| \frac{\kappa \mathcal{P}_t + \gamma \mathcal{N}_t}{100} \right|} \right] \tag{3.3}$$

where $\lambda > 0$, $\kappa \geq 0$, $\gamma \geq 0$ are constant model parameters. From (3.3), the value of f is increasing in $\mathcal{N}_t$ and decreasing in $\mathcal{P}_t$, for fixed values of parameters λ, κ and γ. Furthermore, $f(\mathcal{P}_t \mathcal{N}_t)$ is bounded from above and below by $1/\lambda$ and $1/(\lambda+\max(\gamma,\kappa))$, respectively.

The model structure of the new model is chosen to be one with a direct multiplicative effect of news on the GARCH-predicted volatility:

$$\sigma_t^2 = \left[\frac{1}{\lambda + \left| \frac{\kappa \mathcal{P}_{t-1} + \gamma \mathcal{N}_{t-1}}{100} \right|} \right] (\omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2) \tag{3.4}$$

where $\omega > 0$, $\alpha \geq 0$ and $\beta \geq 0$. In order to keep the unscaled model covariance stationary, the constraint $\alpha + \beta < 1$ is imposed. Note that the stationarity of the scaled model depends on exogenous data (and not simply on the parameters). Furthermore, ε_t is residual returns at time t and defined by:

$$\varepsilon_t = \sigma_t z_t$$

where z_t are standardized residual returns (i.e. i.i.d random variable with zero mean and unit variance), and σ_t^2 is conditional variance. For the sake of brevity, this news augmented GARCH(1,1) model is called as NA1-GARCH(1,1) model.

The chosen model structure adds only three more model parameters and offers a reasonable compromise between increased model complexity and parsimony in terms of model parameters. The choice of model structure is essentially heuristic, and is justified through numerical experiments. One can also choose an alternative model closely related to the above choice in equation (3.4), for instance:

$$\sigma_t^2 = \left[\frac{1}{\lambda + \left(\left| \frac{\kappa \mathcal{P}_{t-1} + \gamma \mathcal{N}_{t-1}}{100} \right| \right)^p} \right] (\omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2) \tag{3.5}$$

where p could be any integer number. However, it is found that using p as a free parameter does not improve results as compared to fixing $p = 1$. Hence the model in equation (3.4) was used throughout the numerical experiments reported here.

The order of the GARCH model has been chosen to be (1, 1) for the NA1-GARCH model; this is firstly due to the fact that from many studies, it can state that this order is sufficient for a good modelling (see, for example, Sharma et al. (1996), Engle et al. (2001) and Baillie and Bollerslev (2002), among others). Secondly taking higher orders could lead to an over-fitting problem. However, results in this study can be easily extended to the case of higher orders. So the proposed model can be seen as a generalization or a completion of many models.

Many other possible model structures have been tested through numerical experiments in order to drive a predictive model to improve the prediction ability of GARCH(1,1) model using news impact scores (positive and negative impact) as proxies of good and bad news stories. The following are some of the tested model structures:

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \kappa\mathcal{P}_{t-1} + \gamma\mathcal{N}_{t-1} \tag{3.6}$$

$$\begin{aligned} \sigma_t^2 &= \omega + \kappa\mathcal{P}_{t-1}\varepsilon_{t-1}^{+} + \gamma\mathcal{N}_{t-1}\varepsilon_{t-1}^{-} + \beta\sigma_{t-1}^2, \\ \text{where} \quad \varepsilon_{t-1}^{+} &= \max(\varepsilon_{t-1}, 0) \text{ and } \varepsilon_{t-1}^{-} = \min(\varepsilon_{t-1}, 0) \end{aligned} \tag{3.7}$$

$$\begin{aligned} \sigma_t^2 = (\omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2)\left(\exp\left(\frac{1}{\lambda + \frac{\kappa\mathcal{P}_{t-1}\varepsilon_{t-1}^{+} + \gamma\mathcal{N}_{t-1}\varepsilon_{t-1}^{-}}{100}}\right)^p\right), \\ \text{where} \quad \varepsilon_{t-1}^{+} = \max(\varepsilon_{t-1}, 0),\ \varepsilon_{t-1}^{-} = \min(\varepsilon_{t-1}, 0) \text{ and } p = 1, 2, \ldots \end{aligned} \tag{3.8}$$

$$\sigma_t^2 = (\omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2)\left(\frac{\kappa\mathcal{P}_{t-1}}{|\gamma\mathcal{N}_{t-1}|}\right)^p, \quad \text{where} \quad p = 1, 2, \ldots \tag{3.9}$$

$$\sigma_t^2 = (\omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2)\left(\frac{|\gamma\mathcal{N}_{t-1}|}{\kappa\mathcal{P}_{t-1}}\right)^p, \quad \text{where} \quad p = 1, 2, \ldots \tag{3.10}$$

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \left(\frac{\kappa\mathcal{P}_{t-1}}{|\gamma\mathcal{N}_{t-1}|}\right)^p, \quad \text{where} \quad p = 1, 2, \ldots \tag{3.11}$$

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \left(\frac{|\gamma\mathcal{N}_{t-1}|}{\kappa\mathcal{P}_{t-1}}\right)^p, \quad \text{where} \quad p = 1, 2, \ldots \tag{3.12}$$

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \Delta, \tag{3.13}$$

$$\text{where} \quad \Delta = \begin{cases} \kappa\mathcal{P}_{t-1} & \text{if } \varepsilon_{t-1} > 0. \\ \gamma\mathcal{N}_{t-1} & \text{if } \varepsilon_{t-1} < 0. \\ 0 & \text{if } \varepsilon_{t-1} = 0. \end{cases}$$

$$\sigma_t^2 = (\omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2)/\Delta, \tag{3.14}$$

$$\text{where} \quad \Delta = \begin{cases} \kappa\mathcal{P}_{t-1} & \text{if } \kappa\mathcal{P}_{t-1} < |\gamma\mathcal{N}_{t-1}|. \\ \gamma\mathcal{N}_{t-1} & \text{if } \kappa\mathcal{P}_{t-1} > |\gamma\mathcal{N}_{t-1}|. \\ 1 & \text{if } \kappa\mathcal{P}_{t-1} = |\gamma\mathcal{N}_{t-1}|. \end{cases}$$

$$(\omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2)(1 + \kappa\sqrt{\mathcal{P}_{t-1}})(1 + \gamma(\mathcal{N}_{t-1}). \tag{3.15}$$

$$(\omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2)(1 + \kappa\sqrt{\mathcal{P}_{t-1}})(1 + \gamma(\mathcal{N}_{t-1})^2). \tag{3.16}$$

$$(\omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2)(1 + \kappa\sqrt{\mathcal{P}_{t-1}})(1 - \gamma(\mathcal{N}_{t-1})^2). \tag{3.17}$$

$$(\omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2)(\eta + \kappa\sqrt{\mathcal{P}_{t-1}})(\eta - \gamma\mathcal{N}_{t-1}), \tag{3.18}$$

where η is constant.

$$(\omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2)(\kappa\sqrt{\mathcal{P}_{t-1}} - \hat{P})(\gamma\mathcal{N}_{t-1} - \hat{N}), \tag{3.19}$$

where $\hat{P}\&\hat{N}$ are very small constants.

$$(\omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2)(1 + \kappa_1\sqrt{\mathcal{P}_{t-1}} - \kappa_2\mathcal{P}_{t-1})(1 + \gamma(\mathcal{N}_{t-1})^2). \tag{3.20}$$

$$(\omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2)(1 + \kappa_1\sqrt{\mathcal{P}_{t-1}} - \kappa_2\mathcal{P}_{t-1})(1 - \gamma(\mathcal{N}_{t-1})^2). \tag{3.21}$$

$$(\omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2)(\frac{1 + \kappa\mathcal{P}_{t-1}}{\eta_1})(\frac{1 - \gamma\mathcal{N}_{t-1}}{\eta_2}), \text{ where } 1.2 \leq \eta_1, \eta_2 \leq 1.8 \tag{3.22}$$

It is also possible to model the news impact as additive, rather than multiplicative (see, for example, equations (3.6), (3.7), (3.11), (3.12) and (3.13)). In this preliminary research, the numerical experiments indicated that an additive news impact model performs a lot worse than a multiplicative news impact model. This is consistent with the intuition that news affects a percentage increase or decrease in volatility, e.g. it is conceivable that a specific negative news will cause x% increase in the *current* level of volatility, rather than causing a specific quantum of increase regardless of the current volatility. Finally, out of those possible model structures stated above in equations from (3.6) to (3.22) the NA1-GARCH model in equation (3.4) provided the best empirical results in initial numerical experiments. Hence NA1-GARCH was chosen for comparison with other GARCH more extensive models in terms of fitting to the data and out of sample performance.

3.4 Model Parameter Estimation

To be able to predict the volatility for a time series, one first has to estimate the parameters of the model from time series data. In this study, the maximum likelihood estimation is applied as a method of estimating the model parameters. An alternative method for estimation would be the generalised method of moments (see e.g. Hall

(2005)); however, MLE is commonly preferred when the distributions are known in closed-form. Four models have been estimated over different datasets for both indices (FTSE100 and S&P500). The models are: simple GARCH(1,1), TGARCH(1,1), EGARCH(1,1) and NA1-GARCH(1,1). In this study, any common parameters are initialised with the same initial values for each of the models. Then the four models are calibrated using the maximum likelihood estimation, each under two different distributional assumptions: Gaussian and Student's t-distribution. MATLAB software was used to calibrate the volatility models and estimate the parameters of the models. The build-in function "*fminsearch*" in MATLAB was used to solve the minimization problem of the negative log-likelihood function.

3.4.1 GARCH model calibration

For a GARCH(1,1) model with Normal conditional returns, the likelihood function is

$$L(\theta|r_1, r_2, \cdots, r_T) = \prod_{i=1}^{T} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(\frac{-(r_i-\mu)^2}{2\sigma_i^2}\right),$$

where θ is a set of the model parameters (here for GARCH model, $\theta = (\omega, \alpha_1, \beta_1, \mu)$). Since the log function is monotonically increasing the function of L, one can maximize the log of the likelihood function

$$\ln L(\theta|r_1, r_2, \cdots, r_T) = -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{i=1}^{T}\ln\sigma_i^2 - \frac{1}{2}\sum_{i=1}^{T}\left(\frac{(r_i-\mu)^2}{\sigma_i^2}\right),$$

where σ_i^2 is given by equation (2.7) for GARCH[1] model. The other three models are estimated in a similar way; for the NA1-GARCH model $\theta = (\omega, \alpha_1, \beta_1, \mu, \kappa, \gamma, \lambda)$ and σ_i^2 is given by equation (3.4); for the EGARCH model $\theta = (\omega, \alpha_1, \beta_1, \gamma_1, \mu)$ and σ_i^2 is given by the exponential of equation (2.13), and $\theta = (\omega, \alpha_1^+, \alpha_1^-, \beta_1, \mu)$ for TGARCH and σ_i^2 is given by the square of equation (2.16).

Tables A.1 and A.2 (*see Appendix A*) show parameter estimations for the four models on all datasets described in subsection (3.2.1).

Notice that, besides estimating the parameters, $\alpha_0, \alpha_1, \beta_1$ and μ, the initial volatility σ_1 has to be estimated. If the time series is long enough, the estimate for σ_1

[1]Please note that for the sake of brevity the order of the model "(1,1)" for all the models are omitted in the rest of this chapter.

will be unimportant.

The error terms of the models given by equations (2.7), (2.13), (2.16) and (3.4) are assumed to have a conditional normal distribution. However, since some financial time series display significant kurtosis, it is also appropriate to investigate and use a conditional Student's t-distribution (see, for example, Bollerslev (1987) and Baillie and Bollerslev (2002)).

3.4.2 GARCH-t model calibration

Student's t-distribution has become a standard benchmark in developing models for asset return distribution because it is able to describe fat tails observed in many empirical distributions. Also, its mathematical properties are well known. In empirical applications, ARCH models are typically estimated by maximum likelihood under the assumption that the errors are conditionally normal distributed. However, in many empirical applications of financial assets, the standardized residuals appear to have fatter tails than the normal distribution. The GARCH-t model of Bollerslev (1987) relaxes the assumption of conditional normality by instead assuming that the standardized innovations follow a standardized Student's t-distribution. Bollerslev proposed a standardized Student's t-distribution with $\nu > 2$ degrees of freedom whose density is given by:

$$f(z_t, \nu) = \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\sigma_t^2(\nu-2)}} \left(1 + \frac{z_t^2}{\nu-2}\right)^{-(\nu+1)/2} \tag{3.23}$$

where $z_t = \varepsilon_t/\sigma_t$ and $\Gamma(\nu) = \int_0^\infty e^{-x}x^{\nu-1}dx$ is the gamma function and $\nu > 2$ is the degrees of freedom and the parameter that measures the tail thickness. The Student's t-distribution is symmetric around mean zero. The log likelihood function for the GARCH-t model is thus given by:

$$\ln L(\theta) = \sum_{t=1}^{T} \ln \left[\Gamma\left(\frac{\nu+1}{2}\right) \Gamma\left(\frac{\nu}{2}\right)^{-1} \left((\nu-2)\sigma_t^2\right)^{-1/2} \left(1 + (\nu-2)^{-1} z_t^2\right)^{-(\nu+1)/2} \right] \tag{3.24}$$

where σ_t^2 is given by equation (2.7) and θ is the vector of parameters to be estimated for the conditional mean, the conditional variance and the density function. When $\nu \to \infty$ the distribution becomes normal, so that the lower ν is, the fatter are the tails. For calibration one has to maximize the log of the likelihood function of equation (3.23).

In this study, the above technique is used to calibrate the four models (GARCH-t, TGARCH-t, EGARCH-t and NA1-GARCH-t) using the maximum likelihood estimation under the Student's t-distribution assumption of the errors distribution. Tables A.3 and A.4 (*see Appendix A*) show parameters estimation for the four models on all datasets described in subsection (3.2.1).

3.5 Methodology and Results

This section presents the analysis and comparison of the volatility estimated by each model in the previous section against observed values of volatility. Four models will be used to compute the volatility for each dataset in both indices (FTSE100 and S&P500). The models are: GARCH, TGARCH, EGARCH and NA1-GARCH. The realized volatility is employed as a volatility benchmark and its values can be calculated by the standard deviation of the asset daily returns. Typically, realised volatility is often measured as the sample standard deviation:

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{t=1}^{N}(r_t - \mu)^2} \tag{3.25}$$

where r_t is the return on day t and μ is the average return over the N-day period. For the purpose of the performance comparison of the models, two criteria will be used. The first criteria is the model fit, and the second criteria is the prediction accuracy of the model.

3.5.1 Model Calibration

In this study the daily frequency data is used for analysis and each dataset contains 750 data points. As mentioned earlier in the previous subsection (3.2.1), 500 data points are used as in-sample period to estimate the parameters of the models, whereas the other 250 data points are kept aside as out-of-sample period for backtesting purposes. Several methods are commonly used for comparing models. Choice of methods depends on whether the models are nested or not. If two models are nested, then the simpler model (the one with fewer parameters) is a special case of the more complicated model (the one with more parameters). If the models are nested, then they can be compared using a likelihood-ratio test. The test statistic is

$$2[-log(L(M_s))] - [-log(L(M_c))] \sim \chi^2(df = P_c - P_s) \tag{3.26}$$

where $L(M_s)$ is the likelihood for the residuals of the simple model, $L(M_c)$ is the likelihood for the residuals of the complicated model, and P_s and P_c are the numbers of parameters in the simple and complicated models, respectively. However, if the models are not nested, then they may be compared using Akaike's information criterion (AIC), which makes adjustments to the likelihood function to account for the number of parameters. AIC can be used for both nested and non-nested models.

The Akaike information criterion, which was proposed in Akaike (1974), is popular among scientists, primarily because it permits comparisons between non-nested models that may differ in terms of number of free parameters. The purpose of the AIC, as explained in Bozdogan (2000), is to select a model that produces the smallest expected discrepancy, where the expectation is taken across the population of replications generated by a fixed design. A complex model may give the smallest discrepancy for the particular replication of the design to which it was fit, but it may give a larger expected discrepancy, averaged across many replications of the design. Multiple models can be compared using AIC whether they are nested or not. Burnham and Anderson (2003) present much more sophisticated methods for evaluating differences in AIC and make a strong case for using AIC in all model comparisons. A closely related method is the Bayesian information criterion (BIC), which was proposed in Schwarz et al. (1978).

The Akaike information criterion and Bayesian information criterion are employed in this study to assess the goodness of fit to the data of the candidate models. However, it is possible that the model with better fit is not good for prediction. According to the AIC and BIC, to evaluate the performance of several models in terms of how well they explain the data the model having lowest AIC or BIC values is considered to have better model fit. The AIC is given by:

$$AIC = -2\ln(L) + 2P, \tag{3.27}$$

and the BIC is given by:

$$BIC = -2\ln(L) + \ln(N)P, \tag{3.28}$$

where L is the maximum value of the likelihood function for the estimated model, P is the number of estimated parameters in the model and N is the number of data points.

In order to evaluate the goodness of fit to the data of the proposed model (NA1-GARCH), the AIC and BIC values of NA1-GARCH(1,1) model are compared with their peers of simple GARCH(1,1) (as a symmetric GARCH-type model), TGARCH(1,1) and EGARCH(1,1) (as asymmetric GARCH-type models). Tables A.9, A.10 in *Appendix* A present the log likelihood (LLH), AIC and BIC values for each estimated model, when the distribution of returns is assumed to be Gaussian. It is clear that in all datasets GARCH and EGARCH models have slightly greater log likelihood values than NA1-GARCH model. This slight difference led the AIC and BIC values of GARCH and EGARCH to be slightly smaller than NA1-GARCH. However, in some of the datasets NA1-GARCH model has greater log likelihood value than TGARCH model. According to the AIC and BIC values of the four models in the tables, it can be clearly seen that in (36 out of 48) cases NA1-GARCH model has smaller AIC and BIC values than TGARCH model. This comparison suggested that GARCH and EGARCH models are a better fit to the data than NA1-GARCH, and NA1-GARCH explains the data better than TGARCH when the returns are assumed to be normally distributed.

Tables A.11, A.12 in *Appendix* A present the LLH, AIC and BIC values for each estimated models, when the distribution of returns is assumed to be Student's t-distribution. GARCH and EGARCH models have slightly greater log likelihood value than NA1-GARCH model in all datasets but the differences in the log likelihood values between them are very small. However, in almost all datasets (47 out of 48) NA1-GARCH model has a greater log likelihood value than the TGARCH model. According to the AIC and BIC values of the four models in the tables, the AIC and BIC values of GARCH and EGARCH models are slightly smaller than NA1-GARCH model in all datasets. Nevertheless, the differences are extremely small (less than 0.5% in average). On the other hand, in (45 out of 48) cases the NA1-GARCH model has smaller value than TGARCH model. As was the case with assuming normal distributed returns, assuming t-distributed returns also suggests that GARCH and EGARCH are a slightly better fit to the data than NA1-GARCH, and NA1-GARCH is a much better fit than TGARCH.

To sum up, according to the log likelihood, AIC and BIC values of the four models, a ranking of estimated models suggest the following ranking from most descriptive to least: GARCH, EGARCH, NA1-GARCH and TGARCH. Keeping in mind that the differences between GARCH and EGARCH, and NA1-GARCH were very small.

3.5.2 Prediction Accuracy

There are many different measures of error which can be used to compare the prediction accuracy of different models. In this study, Mean Absolute Error (MAE) and the Root Mean Squared Error (RMSE) are used as the error criteria. These are defined as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|f_i - y_i| = \frac{1}{n}\sum_{i=1}^{n}|e_i|, \tag{3.29}$$

and

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(f_i - y_i)^2}, \tag{3.30}$$

where n is the number of data points (here is out of sample points), f_i is the one-step ahead predicted volatility and y_i is the realized volatility, which is assumed to be a true value of volatility. Tables A.5, A.6, A.7 and A.8 in *Appendix* A show the MAE and RMSE values of the estimated models: the simple GARCH, TGARCH, EGARCH and NA1-GARCH, each under two different distributional assumptions: Gaussian and Student's t-distribution for each dataset.

After calculating the MAE and RMSE of the estimated models for each dataset when the distribution of the returns is assumed to be Gaussian, a comparison is made between the proposed model (NA1-GARCH) and each of the other three models individually to assess the prediction accuracy for each model. First the MAE values of the NA1-GARCH model and the other models are compared for each dataset. From the Tables A.5 and A.6, it can be seen that in (38 out of 48), (43 out of 48) and (38 out of 48) datasets the MAE values of NA1-GARCH model are less than the MAE values of simple GARCH, TGARCH and EGARCH models, respectively. The RMSE values of NA1-GARCH model are smaller than the RMSE values of simple GARCH, TGARCH and EGARCH models in (37 out of 48), (40 out of 48) and (34 out of 48) datasets, respectively. These results suggest that including news terms in the GARCH model is very likely to improve the prediction of volatility when the distribution of the returns is assumed to be Gaussian.

Tables A.7 and A.8 report the MAE and RMSE values of the four models for each dataset when the distribution of the returns is assumed to be Student's t-distribution. From these two tables, obviously the NA1-GARCH model predicted the volatility better than the other three models. It is clear that the MAE values of NA1-GARCH are smaller than the MAE values of simple GARCH, TGARCH and EGARCH models in (43 out of 48), (47 out of 48) and (43 out of 48) datasets, respectively. It is observed that the NA1-GARCH outperforms the three other models in terms of RMSE values too. In (42 out of 48), (46 out of 48) and (36 out of 48) datasets the RMSE values of NA1-GARCH model are less than the RMSE values of simple GARCH, TGARCH and EGARCH models, respectively. Tables 3.2 and 3.3 give summaries of how many times NA1-GARCH model is better than the other three models in terms of MAE and RMSE for all datasets.

NA1-GARCH model is better than	FTSE100		S&P500	
	Prediction accuracy out of 24 datasets		Prediction accuracy out of 24 datasets	
	in terms of MAE	in terms of RMSE	in terms of MAE	in terms of RMSE
GARCH	20	20	18	17
TGARCH	21	20	22	20
EGARCH	20	17	18	17

Table 3.2: Summary of how many times out of 24, NA1-GARCH is better than GARCH, TGARCH and EGARCH in terms of MAE and RMSE for all datasets when the Gaussian distribution is assumed.

NA1-GARCH model is better than	FTSE100		S&P500	
	Prediction accuracy out of 24 datasets		Prediction accuracy out of 24 datasets	
	in terms of MAE	in terms of RMSE	in terms of MAE	in terms of RMSE
GARCH	22	21	21	21
TGARCH	24	23	23	23
EGARCH	22	17	21	19

Table 3.3: Summary of how many times out of 24, NA1-GARCH is better than GARCH, TGARCH and EGARCH in terms of MAE and RMSE for all datasets when the Student's t-distribution is assumed.

In a nutshell, the MAE and RMSE values of NA1-GARCH model are compared with those of simple GARCH, TGARCH and EGARCH models to assess the prediction accuracy for each model, under two different distributional assumptions: Gaussian and Student's t-distribution for each dataset. The comparison showed that the NA1-GARCH model significantly outperforms the other three models (GARCH, TGARCH and EGARCH) in terms of the prediction power in at least two thirds or more of the datasets as it produced relatively smaller forecasting errors in both MAE and RMSE.

3.6 VaR Backtesting Process

For VaR backtesting, this study uses 95% and 99% confidence levels which are commonly used in the industry. A higher confidence level is not practicable with one year of out-of- sample data and one day time horizon. Throughout the backtesting process daily trading outcomes are compared to daily VaR estimates obtained using the estimated volatility models. Let $\Delta P = P_{t+1} - P_t$, denote profit or loss of the portfolio over one day time interval. The corresponding VaR estimate is then defined as VaR_t, which is calculated at the beginning of the period using the closing prices of the days t and $t+1$. For example, the first VaR estimate is calculated with the closing price of the first day, t, of the out of sample period. This estimate is then compared to the trading outcome (profit or loss) that is realized at the next day, $t+1$. The Variance-Covariance method of one step ahead Value at Risk forecast is

$$VaR_{(t|t-1)} = -\alpha \times \sigma_t, \tag{3.31}$$

where α is 1.645 for 95% confidence level and 2.33 for 99% confidence level and σ_t is the estimated volatility forecasted at time $t-1$ by the candidate model. This forecast assumes the mean of the return distribution to be zero, which is a realistic assumption over a 1-day horizon. Further details can be found in Linsmeier et al. (1996).

This study carried out two different statistical hypothesis tests at both confidence levels and for all the models: Kupiec test and Christoffersen's test. These tests represent a fairly traditional approach to backtesting since they can be applied virtually in every case where VaR figures are computed. Using these tests requires only the number of total observations, and number of VaR violations (or exceptions).

In order to provide some insight into the backtesting process for a variety of risk models (including NA1-GARCH), this study reports results on two different datasets from each index (FTSE100 and S&P500). The choice of two datasets from the 24 data sets discussed in subsection 3.2.1 was driven by the fact that NA1-GARCH model performs differently on these two datasets. For example, in the first dataset the NA1-GARCH model outperforms all the other models in terms of predictive accuracy, whereas it underperforms at least one (or more) of the models in the second dataset. The discussion below (and the results reported in tables A.13 to A.20) are broadly representative of the results on other datasets, which are not reported here. Therefore, as mentioned earlier two different datasets are considered from each index (FTSE100 and S&P500) to assess the validity of NA1-GARCH as a risk model. For each dataset, the models are calibrated under two different distributional assumptions: Gaussian and Student's t-distribution. On the basis of these datasets one can already perform several different backtests; two different backtests have been employed: the unconditional coverage Kupiec test and the conditional coverage Christoffersens's test, which were described earlier in chapter 2. Tables from A.13 to A.20 in *Appendix A* present a summary of results of VaR estimates of the different models with the two datasets. The most important columns in the tables are the observed number of exceptions and the calculated values of the two tests (Kupiec and Christoffersen's test). The tables show 95% and 99% VaR backtesting outcomes of the risk models (GARCH, TGARCH, EGARCH and NA1-GARCH). The out-of-sample data is used to find the observed number of exceptions. The values of the Kupiec test (unconditional coverage) and Christoffersen test (conditional coverage) are calculated using equations (2.29) and (2.30), respectively. Then the computed test statistics are compared with the critical value (3.841) to evaluate the model. The model is considered *not rejected* if the value of the Kupiec test is less than the critical value and the observed number of expectations is less than 7 for 99% confidence interval, and between 6 and 21 for 95% confidence interval as pointed out earlier in the non-rejection regions for Kupiec test table 2.1 in chapter 2. The model is considered *not rejected* if the value of the Christoffersen test is less than the critical value. The last four columns of the tables indicate whether the estimated VaR models, at the specified level of confidence, are rejected or not[2].

The first step is to investigate the datasets where the NA1-GARCH model outperforms the other models. Tables A.13 and A.14 show the backtesting outcomes of four risk models when the models are calibrated on the datasets under the normal distributional assumption. According to the Kupiec test outcomes, most of the estimated VaR models in the first table passed the test (not rejected) at both confidence levels (95% and 99%) except the EGARCH model, which fails to pass the test at confidence level 99% because its Kupiec test value is greater than the critical value of 3.841. The outcomes of Christoffersen's test show that all the estimated VaR models in the first table successfully passed the test at both confidence levels (95% and 99%). From the second table (A.14) it is clear that most of the models passed the Kupiec test at both confidence levels except the TGARCH model, which failed at confidence level of 99% because of the higher value of its Kupiec test. Regarding Christoffersen's test, only the TGARCH model has not passed the test at both confidence levels. Tables

[2]In the tables, the entry "Accepted" stands for the case where the model is not rejected by the corresponding hypothesis test.

A.15 and A.16 show the backtesting outcomes of the four risk models when the models are calibrated under Student's t-distributional assumption. In Table A.15, it can be seen that the models which failed to pass the Kupiec test are the TGARCH model at both confidence levels and the EGARCH model at 99% confidence level, whereas GARCH and NA1-GARCH have passed the Kupiec test at both levels. In addition, all the models were successfully passed the Christoffersen test at both levels. In Table A.16, it is obvious that all the models passed the tests at both confidence levels.

On the other hand, an investigation has been carried out on the dataset when the NA1-GARCH model is worse than at least one or more of the other models. Tables A.17 and A.18 show the backtesting outcomes of the four risk models when the models are calibrated under the normal distributional assumption. In Table A.17, almost all of the estimated VaR models pass the test apart from the NA1-GARCH model, which failed to pass the Kupiec test at confidence level 99%. In the second table (A.18), it is clear that the EGARCH model failed to pass the Kupiec test at both confidence levels and TGARCH model failed to pass the Christoffersen test at both levels, whereas GARCH and NA1-GARCH models passed the tests successfully at both confidence levels. Tables A.19 and A.20 show the backtesting outcomes of the four risk models when the models are calibrated on the datasets under the Student's t-distributional assumption. In Table A.19, it can be seen that all the models passed the Kupiec and Christoffersen tests at both confidence levels except the TGARCH model, which failed to pass the tests at both confidence levels. The reason for the TGARCH model to fail the Kupiec test is because the number of exceptions at both levels has not fallen in the Kupiec test non-rejecting region. Finally, in Table A.20, once again the TGARCH model failed to pass the Kupiec test at both levels and managed only to pass the 95% confidence level of Christoffersen's test. Furthermore, the other three models have successfully passed the tests at both confidence levels.

To sum up, NA1-GARCH cannot be rejected as a suitable VaR model by any of the two tests on the dataset for which it outperforms the other models in terms of out of sample prediction. On the other dataset, it is rejected in one case.

3.7 Conclusion

Numerous studies have suggested that GARCH-type models using asset price data provide good volatility forecasts. The empirical analysis in this chapter demonstrates that GARCH-type models can be enhanced by using exogenous sources of information besides the asset price data to improve the forecasting performance of the models. This study investigates the impact of quantitative news sentiment on the volatility of stock market returns in two different markets. The stock index return volatility of FTSE100 and S&P500 have been modelled and estimated to compare the out of sample predictive ability of GARCH, EGARCH, TGARCH and the proposed NA1-GARCH models under two distributional assumptions: Gaussian and Student's t-distributions. The NA1-GARCH model is a modified version of a GARCH model, which incorporates the impact of news to describe the conditional variance dynamics. The first aim of this study is to examine the impact of news on the index returns volatility and then to find out whether the proposed NA1-GARCH model has a superior forecasting ability compared to the aforementioned models. Furthermore, this study carried out an empirical application to calculating VaR with the different volatility estimates obtained by the above models.

The contribution of this research is to present evidence that including the news impact term in the simple GARCH model improves the prediction power of the model. This study examined a large number of non-linear volatility model structures and proposed a specific model structure on the basis of superior prediction performance. Particular emphasis has been given to the forecasting performance of the models and whether they can capture the characteristics of asset volatility and the influence of news on it. Out of all those possible model structures the NA1-GARCH showed the best empirical results in initial numerical experiments. The empirical results also showed that the prediction performances of the models vary across datasets and error measurement methods. The comparative analysis in evaluating the forecasting performance of the four models (GARCH, TGARCH, EGARCH and NA1-GARCH) shows that news sentiment has an impact on the daily volatility of stock market returns. In addition, the findings reveal that the same magnitude of positive (good news) and negative (bad news) shocks have different impact on the future volatility. The empirical results also show that the Student's t-distribution is more appropriate than the normal assumption, as it generates relatively more accurate forecasts. The empirical analysis suggests that the GARCH and NA1-GARCH models might be more useful than the other two models (TGARCH and EGARCH) when estimating VaR and implementing risk management strategies for FTSE100 and S&P500 stock index returns.

Chapter 4

Forecasting Asset Return Volatility Using Firm-Specific News Data

4.1 Introduction and Background

While the last chapter looked at volatility of stock indices, the objective of this chapter is to look at whether firm-specific news can add value to predicting the volatility of individual stocks. In addition to the research on topics of volatility prediction and the impact of news as reported in the previous two chapters, one can also mention Engle and Ng (1993), who defined the news impact curve that measures how new information is incorporated into volatility estimates. Their results suggest that the EGARCH model, first proposed in Nelson (1991), can capture most of the asymmetry without the need of modelling news impact independently. However, they report evidence that the variability of conditional variance implied by the EGARCH model is too high. The early research of applying news analysis to financial markets focused on equities. Later, macroeconomic news and its impact on fixed income has been studied extensively; see, e.g. Arshanapalli et al. (2006). Li and Engle (1998) studied the effect of the macroeconomic public announcements and revealed the heterogeneous persistence from scheduled news versus non-scheduled news. They examined the reaction of the Treasury futures market to the periodically scheduled announcements of prominent U.S. macroeconomic data.

More recent focus of relevant research is the dynamic relationship between news sentiment and the changes in the asset price dynamics. Ho et al. (2013) compared macroeconomic news sentiment with firm-specific news sentiment and found that the latter accounts for a greater proportion of overall volatility persistence. Crouhy and Rockinger (1997) confirmed that volatility rises more in response to bad news than to good news. Riordan et al. (2013) confirmed that negative news messages induce stronger market reactions than positive ones. The empirical results of Song (2010) indicated that unexpected bad news about a particular portfolio tends to increase the volatility of the returns on other correlated portfolios, whereas unexpected good news about a particular portfolio has an opposite impact on the volatility of correlated portfolios. Chen and Ghysels (2010) found that moderately good (intra-daily) news reduces volatility (the next day), while both very good news (unusual high intra-daily positive returns) and bad news (negative returns) increase volatility, with the latter having a more severe impact. In addition, Sidorov et al. (2013) considered trading volume as a proportional proxy for information arrivals to the market and the daily number of press releases on a stock (news intensity) as an alternative explanatory variable in the basic equation of a GARCH model. They showed that the GARCH(1,1) model augmented with volume does remove GARCH and ARCH effects for the most of the companies, while the GARCH(1,1) model augmented with news intensity has difficulties in removing the impact of log return on volatility. Later Sidorov et al. (2014) analyzed the impact of news intensity as extraneous sources of information on stock volatility. Their results showed that the GARCH(1,1) model augmented with the news intensity performs better than the pure GARCH model.

There is a strong, yet complex relationship between market sentiment and news. Traders and other market participants digest news rapidly and update their asset positions accordingly. However, for models to incorporate news directly and automatically, one require quantitative inputs, whereas raw news is qualitative data. Companies such as RavenPack and Thomson Reuters have developed linguistic analytics which process the textual input of news stories to determine quantitative sentiment scores. Both sets of RavenPack and Thomson Reuters data are similar in structure (see RavenPack (2014) and Reuters (2010)). However, Ravenpack data is used in this study since this was the only data source available from the commercial sponsor (see Mitra and Mitra (2011) for a detailed analysis of RavenPack data). As compared to the amount of effort expended in forecasting volatility from return time series alone, the academic literature on exploiting these sentiment scores for pricing or forecasting seems to be somewhat limited. For example, Tetlock (2007) explored the interactions between investor sentiment and stock market, Mitra et al. (2008) used quantified news and implied volatility to improve risk estimates as the market sentiment and environment changes.

In the work presented here, RavenPack's news sentiment score has been used as a quantitative proxy for news sentiment. This proxy is used as an exogenous term in a modified version of the GARCH model proposed by Bollerslev (1986). A new model structure has been proposed to introduce the impact of news into volatility prediction in a way which is meaningful from an economic point of view. While other GARCH models exist (see, for example, Generalized Error Distribution GARCH (GED-GARCH) by Nelson (1991), Generalized Regime-Switching GARCH (GRS-GARCH) by Gray (1996), the Glosten, Jagannathan and Runkle GARCH (GJR-GARCH) by Glosten et al. (1993) and Semi-parametric Fractional Autoregressive GARCH (SEMIFAR-GARCH) by Feng et al. (2007), among others), I have tested variants of GARCH models, via pure GARCH by Bollerslev (1986), EGARCH by Nelson (1991), TGARCH by Zakoian (1994), GARCH-JUMP by Maheu and McCurdy (2004) and GARCH-t (GARCH with t-distributed residuals) in conjunction with the news proxy. These models were calibrated and tested in terms of volatility forecasting and risk prediction ability, on datasets from two different financial markets. The structure of the proposed model in this study is based on a GARCH(1,1) model. However, the proposed model could be easily extended and modified to account for more general GARCH(p,q) models, with $\max(p, q) > 1$. The simple GARCH(1,1) model has been found to adequately fit many economic and financial time series as well as proven surprisingly successful in predicting conditional variances; see for example, Bollerslev (1987), McCurdy and Morgan (1987); Hsieh (1988), Hsieh (1989), Sharma et al. (1996) and Baillie and Bollerslev (2002). Further, there is evidence in the literature that it is hard to beat GARCH(1,1) in terms of its forecasting ability (see, e.g. Hansen and Lunde (2005)). This was also confirmed in this study in the experiments on news enhanced versions of GARCH, GARCH-t, EGARCH and T-GARCH models. Therefore, the use of higher order GARCH model has not been reported in this study. As mentioned earlier, literature provides some evidence of EGARCH model to capture asymmetry in volatility which may result from differing impact of positive and negative news (see, e.g. Engle and Ng (1993)). Hence these models have been compared with each other as well as the EGARCH(1,1) model on multiple datasets.

The structure of the *second* news augmented GARCH (NA2-GARCH) model is novel and this study vindicates the findings of other researchers, namely, Mitra et al. (2008) and Arbex-Valle et al. (2013) who have used factor models as predictors of realized volatility. The broad conclusion of the earlier studies which is reinforced by this study is that in the financial markets the use of news sentiment leads to better predictions of the volatility of asset returns.

The rest of the chapter is organized as follows. Section 2 presents the granularity of the experimental data and also explains the two streams of time series, namely, the market data and news metadata that are used for numerical experiments. Section 3 describes the model using which the stream of sentiment meta data is turned into impact of news items on asset price: this named as the *news impact* model. Also the new NA2-GARCH model is described. In section 4 the issues of model calibration, model fitting and the performance measures are addressed. In section 5 the computational results of the empirical investigation of the two sets of six assets comparing the performances of the GARCH, NA1-GARCH, NA2-GARCH and EGARCH models using the chosen performance measures is analysed. Section 6 set out the discussions and conclusions.

4.2 Data

In this chapter, the experimental dataset comprises two streams of time series data: the daily market (price) data of the closing prices and the news metadata (supplied by RavenPack) for each asset considered in the experiment. The news meta data provides us quantified values of sentiment which in turn is processed into news impact scores; this news impact model is described in section 4.3.1.

4.2.1 Data Granularity

The trading day starts at 08:00 hours and ends at 16:30 hours. Thus, in a trading day the total number of minutes is 510. The frequency of the news impact scores aligned to the trading hours of 08:00-16:30. Therefore, any news arriving overnight or during the weekend is bucketed into the next morning or days first minute, where the size of a bucket is 1 minute. Hence, the assumption is taken that the impact of such overnight, weekend and holidays news is incorporated into prices the following trading day.

For this study, the stock market daily closing price data has been used. Thus, the news impact score was taken at the last minute in the trading day, which was the aggregation of all the news data from 08:00 to 16:30. The news impact scores actually represented by two data streams and they normalized or scaled in the range between 0 and 1 for positive impact, whereas between -1 and 0 for negative impact. Note that although news data is intraday, it is not conformed to a regular time scale thus there is not a guaranteed number of data points within each day for a particular asset.

4.2.2 Market Data

The time series data used for modelling volatility in this study is the stock market daily closing price of twelve assets from FTSE100 and EUROSTOXX50 (six assets from each index). The data of 12 assets across different sectors based on their cap-weights is extracted. The reason for the concern with the cap-weights is that companies with large market capitalisation have wide coverage of news which guarantees a sufficient number of news data points. The data covers seven different sectors which are Pharmaceuticals & Biotechnology, Insurance, Oil & Gas, Banking, Mobile Telecommunications, Food & Beverage and Chemicals. Table 4.1 lists these sectors and each asset that is belonging to each sector. The data was obtained from Interactive Data.

	FTSE100		**EUROSTOXX50**	
	Asset	**Sector**	**Asset**	**Sector**
1	AstraZeneca	Pharmaceuticals & Biotechnology	Allianz	Insurance
2	Aviva	Insurance	Anheuser-Busch	Food & Beverage
3	BP	Oil & Gas	Banco Santander	Banking
4	GlaxoSmithKline	Pharmaceuticals & Biotechnology	Bayer	Chemical
5	Lloyds Bank	Banking	Deutsche Bank	Banking
6	Vodafone	Mobile Telecommunications	Total	Oil & Gas

Table 4.1: List of 12 assets from FTSE100 and EUROSTOXX50

The data for each asset covers daily closing prices from 3 January 2005 to 31 December 2015 (10 years). A rolling window of size 750, the number of consecutive observations per rolling window, has been chosen and increments of 253 observations between successive rolling windows such that each rolling window starts from the very beginning of January and represents roughly a period of three years. The data for each individual asset is divided into nine datasets such that each dataset contains 750 observations. The data used in model fitting are different from those used in predicting evaluation. Typically, each dataset is divided into two sub-periods. Considering a dataset consists of T observations, $p_1, \ldots, p_T$, the data is divided as $\{p_1, \ldots, p_n\}$ and $\{p_{n+1}, \ldots, p_T\}$ where n is the initial forecast origin. According to Tsay (2008), a reasonable choice for splitting is $n = 2T/3$, where T is the data points and n is the initial forecast origin. Therefore, since each dataset in this research has 750 observations, the first 500 observations (two years) is used as the in-sample data for fitting the models in order to estimate the parameters of the models, while the remaining of 250 observations (one year) are taken as the out-of-sample data and used to evaluate the forecasting performance of the models.

The purpose of choosing 10 years of data is to study the performance of the proposed model (NA2-GARCH) in different economic periods (including a recession period in 2008) and different markets (UK and in Europe).

In this study, daily returns (r_t) were calculated as the continuously compounded returns which are the first difference in logarithm of closing prices of the asset of successive days:

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right) \tag{4.1}$$

where P_t and P_{t-1} are the daily closing price of the index at current and previous days, respectively.

One of the key condition for using GARCH-type models is that the time series data should be stationary. According to Tsay (2005), a time series $\{r_t\}$ is said to be *strictly stationary* if the joint distribution of $(r_{t_1}, \ldots, r_{t_k})$ is identical to that of $(r_{t_1+h}, \ldots, r_{t_k+h})$ for all h, where k is an arbitrary positive integer and $(t_1, \ldots, t_k)$ is a collection of k positive integers. In other words, strict stationarity requires that the joint distribution of $(r_{t_1}, \ldots, r_{t_k})$ is invariant under time shift. This is a very strong condition that is hard to verify empirically. A weaker version of stationarity is often assumed. A time series $\{r_t\}$ is *weakly stationary* if both the mean of r_t and the covariance between r_t and r_{t-l} are time invariant, where l is an arbitrary integer. In applications, weak stationarity enables one to make inferences concerning future observations (e.g. prediction).

In most of the finance literature, it is common to assume that an asset return is weakly stationary. This can be checked empirically provided that a sufficient number of historical returns are available. There are several methods for testing stationarity of a time series, including: unit root tests (e.g. Augmented Dickey-Fuller (ADF) test or Zivot-Andrews test), and a KPSS test (run as a complement to the unit root tests), amongst others. However, it is common to apply an Augmented Dickey-Fuller test (ADF) (see, e.g. Fuller (1976)) to detect the stationarity of the data. The null hypothesis of the augmented Dickey-Fuller test is that a time series exhibits the feature of a unit root process. The alternative hypothesis is that the time series sample is stationary or trend stationary, depending on the specific test used. The ADF statistic, used in the test, is a negative number. The more negative it is, the stronger the rejection of the hypothesis that there is a unit root at some level of confidence. Small p-values in this test (*i.e.* less than 0.05) suggest that the time series is stationary. In this study, we used the ADF test to see whether the time series sample is stationary or not. As expected, all return series from all the datasets from FTSE100 and EUROSTOXX50 exhibit weak stationarity. Hence the time series data can be used for our experiments.

4.2.3 News Metadata

News analytics data is presented in a metadata format where news is given characteristics such as relevance, novelty and sentiment scores according to an individual asset. The analytical process of producing such scores is fully automated from collecting, extracting, aggregating, to categorising and scoring. The result is an individual score assigned to each news article for each characteristic using scales from $0 - 100$. A news sentiment score measures the emotional tone within a news item and varies between positive and negative. Sentiment score is a value falling within a range consisting of a minimum and maximum depicting the overall tone of a news article. Depending on the measurement of scale, the exact polarity of sentiment in the news can be deduced, *i.e.* Thomson Reuters assign probabilities to the moods "Positive", "Neutral" and "Negative" to infer an overall sentiment that is the average of all three scores (Reuters (2010)), whereas RavenPack directly produce a sentiment score belonging to the range 0-100 that then allows a conclusion of positivity or negativity, where 0 is the most negative score and 100 is the most positive score.

As mentioned earlier, the RavenPack news metadata is used in all the experiments in this study. The RavenPack data is represented in two separate files, the first contains equity (company) related analytics and the second one contains global macro analytics. Each record in the equity news analytics file contains 46 fields including a time stamp, reference identifiers, scores for relevance, novelty and sentiment, and unique identifiers for each news story analysed. The explanation of the descriptions and knowledge of RavenPack data fields are represented in the RavenPack news analytics manual (RavenPack (2014)).

4.3 Models

This section introduces two models; these are news impact model and news augmented GARCH model.

4.3.1 News Impact Model

As mentioned in the previous section (4.2.3), the RavenPack's news analytics database is used in this research. In order to derive suitable news impact terms (positive and negative impact scores) and utilize them as proxies of good and bad news, the following two points have to be taken in consideration:

1. An expression has to be found which describes the attenuation of the news sentiment score.

2. The impact of a news item will decay over the time and older news has less effect on the volatility than new news.

To account for these points, the decay model technique has been used that was first proposed by Yu (2014), which reflect the instantaneous impact of news releases and the decay of this impact over a subsequent period of time. The technique combines exponential decay and accumulation of the sentiment score over a given time bucket under observation. In Yu et al. (2015), the decay model is used to construct predictive models for return, volatility and liquidity. The authors modelled asset returns and asset liquidity (in terms of the bid-ask spread) using an AR(2) and AR(3) model, respectively that are extended by the news impact scores. The predictive volatility model was constructed by adding the news impact scores as additional parameters (external regressors) to the variance equation of a GARCH model. Further mathematical details can be found in Yu et al. (2015). The technique can be summarised as follows:

- Let $\{\hat{\mathbf{N}}^a_{\hat{t},d}\}$ be a set of all news events related to the asset a, $\hat{t}$ is time stamp and d is day, where $\hat{t} = 1, \cdots T$ and T is the last time bucket in the day. Each news item in the set is collected from the RavenPack metadata time series.

- Define a new set of filtered news events $\{\mathbf{N}^a_{t,d}\} \subseteq \{\hat{\mathbf{N}}^a_{\hat{t},d}\}$ as accepted under the filter of relevance and novelty scores arriving in the time bucket t of the trading day, where the size of a bucket is one minute and $1 \leq t \leq 510$, since the total number of minutes is 510 in a trading day.

- Map the event sentiment scores (ESS), which belong to the range of [0,100], to scores that belong to the range of $[-50, 50]$ for each filtered news item $\mathbf{N}^a_{t,d}$. These mapped sentiment scores is denoted as $\mathbf{S}(\mathbf{N}^a_{t,d})$.

- The impact of a particular news item exponentially decays over time. The impact score of a news item is modelled as:

$$Impact\ Score = Event\ Sentiment\ Score\ X\ Exponential\ Decay\ Function$$

- For a given asset, all the relevant news items which arrived in the past have an impact on the asset price volatility at the current time bucket t. Therefore, the impact score of an old news item $(\mathbf{N}^a_{\tau,d})$ at current time bucket t would be calculated as:

$$\mathbf{I}_t(\mathbf{N}^a_{\tau,d}) = \underbrace{\mathbf{S}(\mathbf{N}^a_{\tau,d})}_{\text{Sentiment score}} \underbrace{e^{-\lambda(t-\tau)}}_{\text{Decay}}, \qquad t \geq \tau \tag{4.2}$$

  where λ is the exponent which determines the decay rate. The value of λ has to be chosen such that the sentiment value decays to half the initial value in a specific time span.

- Instead of simply aggregating the impact of the news items with positive and negative impact, they are kept separate so that positive and negative effects do not cancel each other. Cancellation reduces the news flow and can lead to misinterpretation.

- Define PIS^a_d and NIS^a_d to be the sets of all the news items with positive and negative impact scores over a specific threshold for asset a on day d, respectively. In particular,

$$PIS^a_d = \left\{ \mathbf{I}_t(\mathbf{N}^a_{t,d}) \mid \mathbf{S}(\mathbf{N}^a_{t,d}) \geq \theta, \forall\ \mathbf{N}^a_{t,d} \right\}$$

  and

$$NIS^a_d = \left\{ \mathbf{I}_t(\mathbf{N}^a_{t,d}) \mid \mathbf{S}(\mathbf{N}^a_{t,d}) \leq -\theta, \forall\ \mathbf{N}^a_{t,d} \right\},$$

  where θ is the threshold expressed as the sentiment value that is considered large enough for inclusion in the impact computation for a given asset. In this study, the value of the the threshold is set as $\theta = 1$.

- To obtain positive and negative news impact scores for a particular day, d, all the positive and negative news impact scores in that day have been aggregated separately for the asset a; such that:

$$\mathbf{P}^a_d = \sum_{t=1}^{t_m} \mathbf{I}_{t_m}(\mathbf{N}^a_{t,d}), \quad \forall \quad \mathbf{I}_{t_m}(\mathbf{N}^a_{t,d}) \in PIS^a_d,$$

and

$$\mathbf{N}_d^a = \sum_{t=1}^{t_m} \mathbf{I}_{t_m}(\mathbf{N}_{t,d}^a), \quad \forall \quad \mathbf{I}_{t_m}(\mathbf{N}_{t,d}^a) \in NIS_d^a,$$

where $t_m = 510$ is the total number of time buckets.

- Finally, in the work presented here, these two time series for positive and negative daily impact scores have been transformed into scaled news impact scores as follows. Define $\mathcal{T} = \{1, 2, \cdots, T\}$ as a finite index set for days over which data is available. Let

$$P_{(1,T)}^a = \max_{t\in\mathcal{T}} \tilde{\mathcal{P}}_t^a, \quad \texttt{and}$$
$$N_{(1,T)}^a = |\min_{t\in\mathcal{T}} \tilde{\mathcal{N}}_t^a|.$$

The daily positive and negative then are defined respectively as:

$$\mathcal{P}_t^a = \frac{\mathbf{P}_t^a}{P_{(1,T)}^a} \quad \text{and}$$
$$\mathcal{N}_t^a = \frac{\mathbf{N}_t^a}{N_{(1,T)}^a}, \tag{4.3}$$

where $t \in \mathcal{T}$. Clearly, $\mathcal{P}_t^a \in [0, 1]$ and $\mathcal{N}_t^a \in [-1, 0]$.

4.3.2 Second News Augmented GARCH Model (NA2-GARCH)

In general, the basic GARCH conditional variance equation (2.5) under normality provides a reasonably good model for analysing financial time series and estimating the conditional volatility. However, in some cases there are aspects of the model which can be improved so that it can better capture the characteristics and dynamics of a particular time series.

Before introducing a new volatility model structure to improve the volatility prediction of a GARCH model using news data, it is worth recalling that trading on financial markets is strongly influenced by public company-specific, macroeconomic or political information flows. As a result, markets react sensitively to news, which is announced on both regular and irregular basis, and news events appear to affect stock return volatility quickly, suggesting that the market incorporates information quickly. The volatility tends to be higher in a falling market than in a rising market. Therefore, companies such as RavenPack and Thomson Reuters have started to automatically track and monitor relevant information on ten of thousand of companies and quantifying the content of news articles about them, and allowing for the measurement of the reaction to positive, neutral and negative news.

In this study, the aim is to develop a volatility prediction model in which enables one to use the whole information content of news events in order to improve the existing volatility prediction GARCH (1,1) model. Therefore, in the first stage the news impact scores are constructed, as derived in section (4.3.1), so as to be used as proxies of news events in the new model. To accomplish this, RavenPack's news analytics database has been used and its quantitative sentiment scores are exploited. Since the financial markets are mainly sensitive to good and bad news, thus only positive and negative news sentiments are taken into consideration as it is believed that neutral news does not have any affect on stock return volatility.

The model structure needs to reflect the following economic realities: positive and negative news impact the volatility differently. Furthermore, positive news tends to reduce volatility whereas negative news tends to increase volatility. Finally, the impact of news on return's volatility of an asset decays relatively slowly (more as a power law than as an exponential).

To model this effect of news in addition to serial correlation, a scaled version of GARCH model is defined, where the scaling factor is determined by the news sentiment score in the following way. Consider a function of two variables x and y:

$$f(x, y) = a + 0.5 * b \left(\frac{e^x - 1}{e^x + 1} - \frac{e^y - 1}{e^y + 1} \right), \tag{4.4}$$

where a and b are constants. It is easy to see that $f(x, y)$ lies between $(a, a + b)$ for any non-negative values of x and non-positive values of y.

Let $\{\mathcal{P}_t\}$ and $\{\mathcal{N}_t\}$ be two different time series as defined in equation (4.3)[1]. Keeping in mind the economic realities mentioned above, the following function is defined as a scaling factor to GARCH model:

$$f(\mathcal{P}_t, \mathcal{N}_t) = a + 0.5 * b \left(\frac{e^{\kappa \mathcal{P}_t} - 1}{e^{\kappa \mathcal{P}_t} + 1} - \frac{e^{\gamma \mathcal{N}_t} - 1}{e^{\gamma \mathcal{N}_t} + 1} \right), \tag{4.5}$$

where 0.5 is a scaling factor of the function, and a, b, κ and γ are parameters of the model. For example, if the parameters value are set as $a = 0.8$, $b = 0.8$, $\kappa = 4$ and $\gamma = 4$, the outcome will be as illustrated in Figure 4.1. In this example, it can be

[1]The superscript for asset is omitted for simplicity since only one asset is considered at a time.

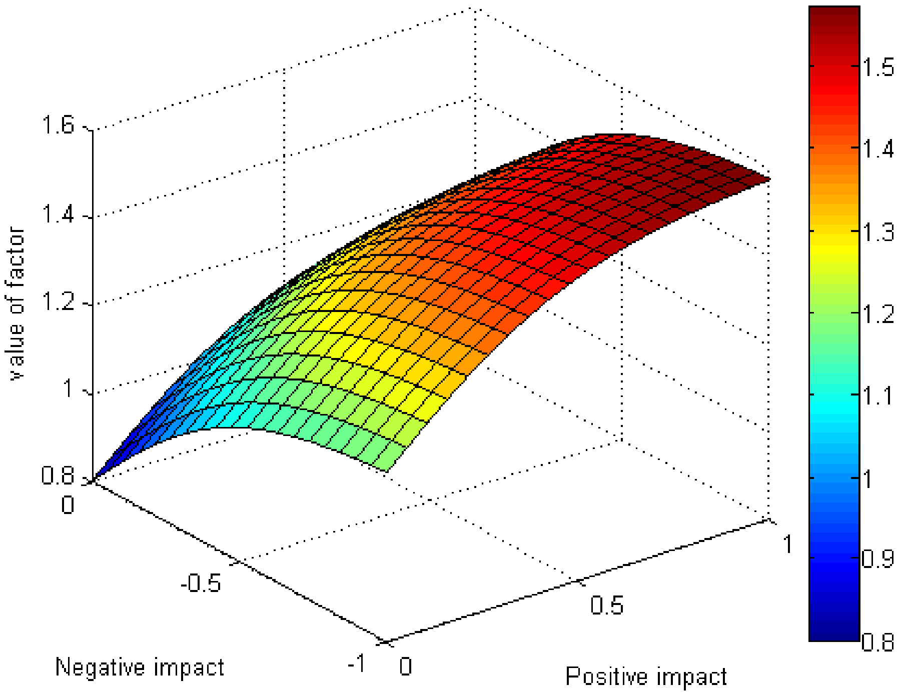

Figure 4.1: Plot of the scaling factor model against positive and negative impact scores

clearly seen that the function reaches its highest value only when $\mathcal{P}_t = 1$ and $\mathcal{N}_t = -1$ (their highest scores). This can be interpreted in our model that the volatility increases when the value of news impact scores ($\mathcal{P}_t$ and $\mathcal{N}_t$) are increasing and vice versa.

The model structure of the *second* news augmented GARCH(1,1) model is chosen to be one with a direct multiplicative effect of news on the GARCH-predicted volatility:

$$\sigma_t^2 = \left[a + 0.5 * b \left(\frac{e^{\kappa \mathcal{P}_{t-1}} - 1}{e^{\kappa \mathcal{P}_{t-1}} + 1} - \frac{e^{\gamma \mathcal{N}_{t-1}} - 1}{e^{\gamma \mathcal{N}_{t-1}} + 1} \right) \right] (\omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2), \tag{4.6}$$

where $a > 0,\ b > 0,\ \kappa \geq 0,\ \gamma \geq 0,\ \omega > 0,\ \alpha \geq 0$ and $\beta \geq 0$. As in the NA1-GARCH case, we impose the stationarity constraint $\alpha + \beta < 1$ only on the unscaled model, as the scaled coefficients depend on exogenous data. In addition, also a constraint on the upper bound of the interval range is needed that is specified by summing up the parameters a and b to keep the news impact related scaling in a reasonable range.

The choice used here is $0.5 \leq a + b \leq 2$, *i.e.* the impact of news is assumed to change the GARCH prediction at most by a factor of 2. Furthermore, ε_t is residual returns at time t and defined by:

$$\varepsilon_t = \sigma_t z_t,$$

where z_t is standardized residual returns (*i.e.* i.i.d random variable with zero mean and unit variance), and σ_t^2 is conditional variance. For the sake of brevity, from now on the *second* news augmented GARCH(1,1) model will be called as NA2-GARCH model.

It is also possible to model the news impact as additive, rather than multiplicative. However, the numerical experiments in this study indicated that an additive news impact model performs a lot worse than a multiplicative news impact model. This is consistent with the intuition that news affects a percentage increase or decrease in volatility, e.g. it is conceivable that a specific negative news will cause x% increase in the *current* level of volatility, rather than causing a specific quantum of increase regardless of the current volatility.

The chosen model structure adds only four more model parameters for each asset and offers a reasonable compromise between increased model complexity and parsimony in terms of model parameters. The choice of model structure is essentially heuristic, and is justified through numerical experiments. One can also reduce the number of parameters by keeping the value of a and b fixed, for instance, $a = 0.5$ and $b = 1.5$. However, treating a and b as free parameters does improve results in terms of predictive ability of the model.

Finally, as mentioned earlier in section 4.1, the model structure of NA2-GARCH model can easily be extended to NA2-GARCH(p,q) model. We chose not to do so due to extensive evidence in the literature on the adequacy of GARCH(1,1) model for forecasting, as mentioned earlier. GARCH(1,1) model has been found to adequately fit many economic and financial time series as well as proven surprisingly successful in predicting conditional variances. Further, there is evidence in the literature that it is hard to beat GARCH(1,1) in terms of its forecasting ability and this was also confirmed in this empirical study.

4.4 Methodology: Calibration and Performance

This section explains the methods that are used in this study for parameter estimation of the models and for the analysis of the datasets of the assets. To compare the performance of the models, two criteria are used. The first criteria is the model fit, and the second one is the prediction accuracy of the model.

4.4.1 Parameter Estimation and Model Fitting

To be able to predict the volatility for a time series, one has to fit the model to the time series. This is done via estimation of the parameters in the studied models. The most common method of this estimation has been used in this study which is the maximum-likelihood estimation (MLE), as described in section 2.2.2. Four models have been estimated over different datasets for all assets in both indices: simple GARCH (described in section 2.1.2), NA1-GARCH (described in section 3.3), NA2-GARCH (described in section 4.3.2) and EGARCH (described in section 2.1.4). In this study, the estimated parameters are initiated to the same initial assumptions for all the models. Then the parameters of the four models have been estimated by using the maximum likelihood estimation, each under Gaussian distributed assumptions. To calibrate a model, maximum likelihood estimation requires one to maximize the likelihood function $L(\theta)$ with respect to the unknown parameter θ. For instance, the likelihood function of a simple GARCH(1,1) model with normal conditional returns is

$$L(\theta|r_1, r_2, \cdots, r_T) = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(\frac{-(r_t-\mu)^2}{2\sigma_t^2}\right), \tag{4.7}$$

where θ is a set of the model parameters (i.e. $\theta = (\omega, \alpha_1, \beta_1, \mu)$). Since the logarithm is monotonically increasing the function of L, then it's equivalent to minimizing the negative log of the likelihood function to estimate the unknown parameters

$$\ln L(\theta|r_1, r_2, \cdots, r_T) = -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\ln\sigma_t^2 - \frac{1}{2}\sum_{t=1}^{T}\left(\frac{(r_t-\mu)^2}{\sigma_t^2}\right), \tag{4.8}$$

where σ_t^2 is given by equation (2.7) for GARCH model. For the NA2-GARCH model $\theta = (\omega, \alpha_1, \beta_1, \mu, \kappa, \gamma, a, b)$ and σ_t^2 is given by equation (4.6), and for the EGARCH model $\theta = (\omega, \alpha_1, \beta_1, \gamma_1, \mu)$ and σ_t^2 is given by the exponential of equation (2.13).

In this study, the low frequency data is used for analysis with 750 daily returns as data points for each dataset. As mentioned in section 4.2.2, the first 500 data points are used as in-sample period to estimate the parameters of the model, whereas the other 250 data points are kept aside as out-of-sample period for backtesting purposes. For the purpose of *in-sample* comparison of the models, the Akaike information criterion (AIC) has been used to compare the goodness of fit to the data of the models. One can calculate the AIC of a model using the following formula:

$$AIC = -2ln(L) + 2P, \tag{4.9}$$

where L is the maximum value of the likelihood function for the model and P is the number of estimated parameters in the model (see, for example, Burnham and Anderson (2003)). While AIC tells us about in-sample performance of the model, out-of-sample performance is often far more important from a forecasting point of view. The measures of out-of-sample performance are considered in the next subsection.

4.4.2 Performance Measures

There are many statistical methods which can be used to observe the prediction accuracy of a model, for instance, Mean Absolute Error (MAE), Root Mean Square Errors (RMSE) and Mean Average Percentage Error (MAPE). In the quantitative comparison between two models in terms of prediction accuracy, the model which produces the smallest values of the forecast evaluation statistics is judged to be the best model. For this study, the mean absolute error and root mean square errors have been considered as the criteria for prediction accuracy since the squared daily returns may not be a proper measure to assess the forecasting performance of the different GARCH models for conditional variance (see Andersen and Bollerslev (1998)). The mean absolute error is given by

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|f_i - y_i| = \frac{1}{n}\sum_{i=1}^{n}|e_i|. \tag{4.10}$$

The root mean squared error is given by

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(f_i - y_i)^2}. \tag{4.11}$$

In both equations (4.10) and (4.11), f_i is the predicted value and y_i is the realized volatility, which is considered as a true value of volatility. Since the errors are squared before they are averaged, the RMSE gives a relatively high weight to large errors.

4.5 Computational Results

In this section, the results of parameter estimation and volatility forecasting for four models are analysed: GARCH (as described in section 2.1.2), NA1-GARCH (as described in section 3.3), NA2-GARCH (as described in section 4.3.2) and EGARCH (as described in section 2.1.4). These results are compared against the observed values of realized volatility. The realized volatility is considered as a volatility benchmark and its values can be calculated by the standard deviation of the asset daily returns. Typically, realised volatility is often measured as the sample standard deviation:

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{t=1}^{N}(r_t - \mu)^2} \tag{4.12}$$

where r_t is the return on day t and μ is the average return over the N-day period.

As mentioned earlier in section (4.2.3), the RavenPack's news analytics database has been employed and some quantitative scores from its fields such as relevance, novelty and sentiment scores are exploited in this research. The event sentiment score is a relative number which describes the degree of positivity and negativity in a piece of news. During the trading day, as news arrives it is given a sentiment value. In order to derive a suitable news impact time series (positive and negative impact scores) and to utilize them as proxies of good and bad news, first the news metadata is filtered so as for the chosen assets, from FTSE100 and EUROSTOXX50, were selected under the filter of relevance score of 100, and any news item that had an event novelty scores under the value of 70 was ignored and not included in the dataset. Then, the ESSs have been transformed into a scaled sentiment score in the range +50 to -50, as it has been found that such a derived single score provides a relatively better interpretation of the mood of the news item. As mentioned in Yu (2014), the value of λ in equation (4.2) has to be chosen such that the sentiment value decays to half the initial value in a specific time span. Thus, in this study the value of λ, which is the exponent that determines the decay rate, is set to 0.007701635 that reduces the initial sentiment value to half in 90 minutes. Finally the steps described in section (4.3.1) are followed to generate two time series of positive and negative news impact scores and use them in NA2-GARCH(1,1) model in equation (4.6). Obviously, other decay functions exist.

One popular choice is the Hill decay function (see Hill (1910)), which is defined as

$$f(t, \tau, \lambda) = \frac{1}{\left(1 + \frac{(t-\tau)}{\lambda}\right)^n} \qquad t \geq \tau \tag{4.13}$$

where t is current time, λ is the chosen value which make the sentiment value decays to half of its initial value (half-life) in a specific time span and n is the hill coefficient. The value of λ is fixed to 0.007701635 similar to the one used in the exponential decay function to generate the news impact scores. Then the exponential decay function is compared with the Hill decay function with the Hill coefficient n=2. It is found that, in all the datasets under study, the new proposed model NA2-GARCH performs better in terms of volatility forecasting when an exponential decay function is used for generating the news impact scores, as compared to the performance when a Hill decay function is used for generating the same scores. Tables 4.2 and 4.3 show the differences in terms of MAE and RMSE between NA2-GARCH with news impact scores that used exponential function and NA2-GARCH with news impact scores that used Hill function as a decay function for all chosen assets from FTSE100 and EUROSTOXX50. The entries in Table 4.2 can be interpreted as follows: for the first asset, "AstraZeneca", it can be seen that the new volatility model (NA2-GARCH) predicted the return volatility better in nine out of nine datasets, in terms of MAE and RMSE, using the news impact scores that were generated by using the exponential decay function in the method described in section (4.3.1), whereas the model predicted the return volatility better only in three out of nine datasets using the news impact scores that were generated by using the Hill function. It is possible that using a function with a slower decay of news impact (such as the power law decay in Hill function) might be more appropriate when the news is infrequent. In all datasets used in this study, the assets are very liquid, news is frequent and correspondingly the impact of a particular news item can be assumed to die out faster than a power law. Hence the use of a Hill decay function is not considered further in this study and the exponential decay function is employed only. The methodology, of course, is applicable to other decay functions.

After computing the news impact scores and calculating the returns for each asset, the four models (GARCH, NA1-GARCG, NA2-GARCH and EGARCH) have been calibrated using maximum likelihood under the assumption of a Gaussian error distribution. The expression for σ^2 in each model is substituted in the normal maximum likelihood and then maximized with respect to the parameters. The *fminsearch* function in MATLAB software is used in the empirical experiments to estimate the parameters.

	Assets	Exponential		Hill	
		MAE	**RMSE**	**MAE**	**RMSE**
1	AstraZeneca	**9**	**9**	3	3
2	Aviva	**7**	**7**	5	6
3	BP	**7**	**7**	4	3
4	GlaxoSmithKline	**8**	**9**	6	7
5	Lloyds Bank	**7**	**7**	1	2
6	Vodafone	**8**	**8**	6	5

Table 4.2: Contingency table for NA2-GARCH model shows the performance differences of the model using different decay functions (Exponential and Hill) in terms of MAE and RMSE for the chosen FTSE100 assets. The greater of the successful cases are highlighted in boldface.

	Assets	Exponential		Hill	
		MAE	**RMSE**	**MAE**	**RMSE**
1	Allianz	**7**	**8**	5	5
2	Anheuser-Busch	**7**	**7**	6	5
3	Banco Santander	**8**	**8**	4	2
4	Bayer	**9**	**8**	4	4
5	Deutsche Bank	**8**	**8**	4	5
6	Total	**7**	**6**	4	3

Table 4.3: Contingency table for NA2-GARCH model shows the performance differences of the model using different decay functions (Exponential and Hill) in terms of MAE and RMSE for the chosen EUROSTOXX50 assets. The greater of the successful cases are highlighted in boldface.

The computational results of the empirical experiments in this study show that the sets of coefficient estimates are quite different from one dataset to another. Tables (from B.1 To B.12) in the *Appendix* B show the parameter estimates of the models for the returns of 12 assets from FTSE100 and EUROSTOXX50 indices. From these tables, the parameters α and β for GARCH, NA1-GARCH and NA2-GARCH models are consistent and their sums are less than unity implying that the models are stationary, though the volatility is fairly persistent since $(\alpha + \beta)$ is close to one. In order to see the goodness of fit to the data of the models, the model fit of NA2-GARCH(1,1) is compared with simple GARCH(1,1), NA1-GARCH(1,1) and EGARCH(1,1) when the distribution of returns are assumed to be Gaussian.

To compare the four models in terms of goodness of fit to the data, the Akaike information criterion (AIC) is applied and the Log-Likelihood values are calculated. The results show that the GARCH and EGARCH models have slightly greater log-likelihood value than NA2-GARCH, which leads to the values of AIC for GARCH and EGARCH models to be slightly smaller than NA2-GARCH. On the other hand, NA1-GARCH and NA2-GARCH models have approximately equal values of log-likelihood and AIC. In most of the cases GARCH and EGARCH have the same AIC values, and the differences between them and the NA2-GARCH model in terms of AIC values are extremely small (less than 1% in average), see tables (B.13 - B.24) in the *Appendix*. Therefore, according to the AIC values of the four models, it is noticed that all of the models have almost the same level of goodness of fit.

To test the predictive ability of the models, the proposed model NA2-GARCH is compared to the simple GARCH model, NA1-GARCH model and the EGARCH model for each dataset by computing the MAE and RMSE values, when the distribution of the returns is normal. Tables (from B.13 to B.24) in the *Appendix* show the MAE and RMSE results of the models for the returns of 12 assets from FTSE100 and EUROSTOXX50. In addition, if the new model NA2-GARCH correctly incorporates the understanding of the impacts of news, it should be able to generate similar patterns as realized volatility in the forecasted volatility. Table 4.4 shows how many times the proposed model (NA2-GARCH) was better than GARCH in terms of MAE and RMSE for each of the chosen assets in FTSE100. The construction of the table is as follows: there are three main columns in the table with headings "Assets" , "MAE" and "RMSE". The chosen assets are listed in the first column under the heading "Assets", the second column with the heading "MAE" is further divided into four sub-columns each labelled or named with a number (from six to nine) that indicates the number of cases in which NA2-GARCH model leads to a lower error than the GARCH model, out of the nine datasets for each asset. For example, if the first sub-column, "6" , is ticked for the first row, "AstraZeneca" asset, then this means our new volatility model (NA2-GARCH) has predicted the volatility better than GARCH model in six out of nine datasets in terms of MAE computations for that particular asset, and so on for the rest of sub-columns. The third column (RMSE) is also divided into four sub-columns, and has the same construction as the second column, but it is showing the number of successful cases of NA2-GARCH based on the RMSE calculations. From this table, it can be seen that the new volatility model (NA2-GARCH) predicted the return volatility better than a GARCH model for the first asset, "AstraZeneca", in every datasets in terms of MAE and RSME; since all the related sub-columns in the first row are ticked. This means that the NA2-GARCH model is 100% a better choice than a GARCH model to predict the volatility for this asset, whereas the NA2-GARCH model predicted the volatility better than the GARCH model only in seven out of nine datasets in terms of MAE and RMSE for the second asset, "Aviva".

Assets		**MAE**				**RMSE**			
		6	**7**	**8**	**9**	**6**	**7**	**8**	**9**
1	AstraZeneca	✓	✓	✓	✓	✓	✓	✓	✓
2	Aviva	✓	✓			✓	✓		
3	BP	✓	✓			✓	✓		
4	GlaxoSmithKline	✓	✓	✓		✓	✓	✓	✓
5	Lloyds Bank	✓	✓			✓	✓		
6	Vodafone	✓	✓	✓		✓	✓	✓	
Total successful cases		6	6	3	1	6	6	3	2

Table 4.4: Contingency table for MAE and RMSE shows the successful cases of NA2-GARCH. The sub-columns represent how many times NA2-GARCH model performed better than GARCH model, and the rows represent the chosen FTSE100 Assets

The prediction ability of the models for the other six assets from EUROSTOXX50 have been examined as well. Table 4.5 shows how many times our new model (NA2-GARCH) was better than GARCH model in terms of MAE and RMSE for each of the chosen assets in EUROSTOXX50.

Tables 4.6 and 4.7 show how many times the NA2-GARCH model was able to

Assets		**MAE**				**RMSE**			
		6	**7**	**8**	**9**	**6**	**7**	**8**	**9**
1	Allianz	✓	✓			✓	✓	✓	
2	Anheuser-Busch	✓	✓			✓	✓		
3	Banco Santander	✓	✓	✓		✓	✓	✓	
4	Bayer	✓	✓	✓	✓	✓	✓	✓	
5	Deutsche Bank	✓	✓	✓		✓	✓	✓	
6	Total	✓	✓			✓			
Total successful cases		6	6	3	1	6	5	4	0

Table 4.5: Contingency table for MAE and RMSE shows the successful cases of NA2-GARCH. The sub-columns represent how many times NA2-GARCH model performed better than GARCH model, and the rows represent the chosen EUROSTOXX50 Assets

predict the volatility better than the EGARCH model in terms of MAE and RMSE for each of the chosen assets in FTSE100 and EUROSTOXX50, respectively. All tables have the same construction. Furthermore, the results show that the NA2-GARCH model correctly incorporated the understanding of the impacts of news as it was able to generate visually very similar plots as realized volatility in the out of sample periods for all datasets.

To sum up, from the Tables 4.4 and 4.5 one can clearly see that NA2-GARCH performs better than a GARCH model in terms of prediction accuracy in at least two thirds (six out of nine) or more of the datasets for each asset. Overall, there are 216 comparisons, with nine datasets, 12 assets and two error metrics. Out of these, the NA2-GARCH model turns out to be a better model than the GARCH model in 184 comparisons out of 216. Further, it is obvious from the Tables 4.6 and 4.7 that NA2-GARCH outperforms the EGARCH model in terms of prediction accuracy in at least two thirds (seven out of nine) or more of the datasets for each asset. Again, out of 216 comparisons, NA2-GARCH model turns out to be a better model than the EGARCH model in 185 comparisons. Tables 4.8 and 4.9 show that NA2-GARCH performs slightly better than NA1-GARCH model in terms of prediction accuracy in at least 5 or more out of 9 datasets for each asset. Therefore, NA2-GARCH performed slightly better than NA1-GARCH model in 114 out of 216 comparisons (i.e. more than half). Furthermore,

	Assets	**MAE**				**RMSE**			
		6	**7**	**8**	**9**	**6**	**7**	**8**	**9**
1	AstraZeneca	✓	✓			✓	✓	✓	
2	Aviva	✓	✓	✓		✓	✓	✓	
3	BP	✓	✓			✓	✓		
4	GlaxoSmithKline	✓	✓			✓	✓	✓	
5	Lloyds Bank	✓	✓	✓		✓	✓	✓	
6	Vodafone	✓	✓			✓	✓		
Total successful cases		6	6	2	0	6	6	4	0

Table 4.6: Contingency table for MAE and RMSE shows the successful cases of NA2-GARCH. The sub-columns represent how many times NA2-GARCH model performed better than EGARCH model, and the rows represent the chosen FTSE100 Assets

	Assets	**MAE**				**RMSE**			
		6	**7**	**8**	**9**	**6**	**7**	**8**	**9**
1	Allianz	✓	✓	✓		✓	✓		
2	Anheuser-Busch	✓	✓			✓	✓		
3	Banco Santander	✓	✓	✓		✓	✓	✓	
4	Bayer	✓	✓	✓	✓	✓	✓	✓	✓
5	Deutsche Bank	✓	✓	✓		✓	✓	✓	
6	Total	✓	✓	✓		✓	✓	✓	
Total successful cases		6	6	5	1	6	6	4	1

Table 4.7: Contingency table for MAE and RMSE shows the successful cases of NA2-GARCH. The sub-columns represent how many times NA2-GARCH model performed better than EGARCH model, and the rows represent the chosen EUROSTOXX50 Assets

Assets		**MAE**				**RMSE**			
		5	**6**	**7**	**8**	**5**	**6**	**7**	**8**
1	AstraZeneca	✓	✓			✓	✓		
2	Aviva	✓				✓			
3	BP	✓				✓	✓		
4	GlaxoSmithKline	✓	✓			✓	✓		
5	Lloyds Bank	✓	✓			✓			
6	Vodafone	✓				✓			
Total successful cases		6	3	0	0	6	3	0	0

Table 4.8: Contingency table for MAE and RMSE shows the successful cases of NA2-GARCH. The sub-columns represent how many times NA2-GARCH model performed better than NA1-GARCH model, and the rows represent the chosen FTSE100 Assets

Assets		**MAE**				**RMSE**			
		5	**6**	**7**	**8**	**5**	**6**	**7**	**8**
1	Allianz	✓				✓	✓		
2	Anheuser-Busch	✓				✓			
3	Banco Santander	✓				✓			
4	Bayer	✓	✓			✓	✓		
5	Deutsche Bank	✓	✓			✓	✓	✓	
6	Total	✓	✓			✓	✓		
Total successful cases		6	3	0	0	6	4	1	0

Table 4.9: Contingency table for MAE and RMSE shows the successful cases of NA2-GARCH. The sub-columns represent how many times NA2-GARCH model performed better than NA1-GARCH model, and the rows represent the chosen EUROSTOXX50 Assets

all the sample means were positive, for instance, the sample means of differences $(MAE_{GARCH} - MAE_{NA2-GARCH})$ and $(RMSE_{GARCH} - RMSE_{NA2-GARCH})$ are both positive, for both indices FTSE100 and EUROSTOXX50. This analysis suggests that including the news impact term in the GARCH framework has improved the predictive ability of GARCH model.

4.6 Discussion and Conclusion

Forecasting accurately future volatility and correlations of financial asset returns is essential since volatility can be used on its own such as in the hedge fund portfolio, or it can be used in conjunction with return measures such as in Sharpe and Sortino ratio formulae. Therefore, the flexibility of a GARCH model and its forecasting accuracy, place the model in a unique position to achieve many of the requirements of the practitioners. However, its use is restricted to one time series data (market data) and ignoring other types of data, especially the market news sentiment that can be very helpful to anticipate the potential impacts on the return volatility of an asset.

This research investigates the impact of high-frequency public news sentiment on the daily log returns volatility for 12 assets from FTSE100 and EUROSTOXX50; a time period from 3 January 2005 to 31 December 2015. A new volatility model has been proposed, namely News Augmented GARCH (NA2-GARCH) model, which enables us to use the news sentiment score to improve the predictive ability of a GARCH model. The computational results of the empirical investigation of the two sets of six assets are analysed, comparing the performances of the GARCH, NA1-GARCH, NA2-GARCH and EGARCH models using the chosen performance measures. The results of the empirical experiments in this research clearly demonstrate that NA2-GARCH provides a superior prediction of volatility than the simple GARCH, NA1-GARCH and EGARCH models. The superior performance of the NA2-GARCH model as compared to the NA1-GARCH model may be attributed to a more flexible model structure. Relative to the GARCH model, NA2-GARCH model has four extra parameters (a, b, κ and γ), with the squared volatility amplification factor ($f(\mathcal{P}_t\mathcal{N}_t)$ in the equation (4.6)) being bounded from above and below by $a + b$ and a respectively. On the other hand, NA1-GARCH model has only three extra parameters (λ, κ and γ) and the scaling function ($f(\mathcal{P}_t\mathcal{N}_t)$ in the equation (3.4)) is bounded from above and below by $1/\lambda$ and $1/(\lambda+\max(\gamma,\kappa))$, respectively. Thus NA2-GARCH model has parametric upper and lower bounds which are independent of the two parameters multiplying

$\{\mathcal{P}_t\}$ and $\{\mathcal{N}_t\}$. This fact might contribute to its outperformance over NA1-GARCH when modelling the volatility of the individual assets.

The empirical results in this study suggest that, when news sentiment score is available, NA2-GARCH is a far better choice than GARCH and EGARCH for volatility prediction. In particular, the NA2-GARCH model incorporates the primary characteristics of historical return volatility clustering with news sentiment scores, which make it capable of capturing the more general features of volatility and provide more robust and transparent predictive abilities over longer, out of sample time horizons. In this study, pre-processed news sentiment data from the commercial provider *RavenPack* has been used. It is reasonable to expect that a volatility prediction model and methodology developed in this research, which works with news sentiment score from one data provider, would also work with score from another data provider. Furthermore, the findings of this study also show that the positive news tends to reduce volatility whereas negative news tends to increase volatility, which is consistent with the studies by Crouhy and Rockinger (1997), Song (2010) and Chen and Ghysels (2010) about the effect of positive and negative shocks. This analysis suggests that including the news impact term in the GARCH framework has improved the predictive ability of GARCH model. Another suggestion is that the use of an exponential decay function is good when the news flow is frequent, whereas the Hill decay function is good only when there are scheduled announcements, because the impact of the news on the market dies slowly as the impact of new information on a stock market depends on how unexpected the news is. NA2-GARCH is thus a computationally efficient means of exploiting the news sentiment score for better volatility prediction and it has the potential to be very useful in industrial practice. Therefore, the findings are crucial for all investors who are trading on the variance or volatility swaps, and can be used to speculate on future realized volatility, or to hedge the volatility exposure of other positions since the profit and loss from a variance swap depends directly on the difference between realized and implied volatility of a given underlying asset. In addition, the NA2-GARCH model would be useful for investors who are focusing on risk-adjusted returns, especially those that utilize asset allocation and volatility targeting strategies. Furthermore, the NA2-GARCH model can be used to estimate VaR more accurately.

Chapter 5

Forecasting Crude Oil Futures Prices Using Macroeconomic News Data

5.1 Introduction and Background

Commodity price and its price fluctuation both play crucial roles in affecting the global economy, even though the impacts are irregular depending on different factors such as geographic regions, specific commodity sectors, time periods etc. In particular, crude oil is an important commodity and a necessary component for the economic development and growth for industrialized and developing countries. Crude oil spot and futures contract prices can significantly impact inflation, unemployment rate, trade, poverty, and other economic conditions in many countries. Crude oil futures are exchange-traded contracts in which the contract is an agreement to buy or sell a specific quantity of crude oil at a predetermined price, on a future delivery date. The impact of crude oil price variation reaches a large number of goods and services. Modelling the variation in price of spot and futures prices of crude oil is important for a variety of market participants, from sovereign governments to futures traders. Central banks and private sector forecasters consider the crude oil price and its futures prices as key variables in generating macroeconomic scenarios. Airlines invest in oil futures in order to hedge their exposure to aviation fuel, as fuel represents a large proportion of their cost. Given the overall importance of crude oil to economy, tools to improve the quality of prediction of oil prices and oil futures prices are clearly desirable.

The aforementioned need for accurate forecasting of oil spot and futures prices have attracted a lot of academic research. The relation between futures prices and spot price has been the centre of attention for a large number of studies, and the literature is rich with several studies covering a range of aspects with respect to this relationship. As a result, the literature on using futures prices for forecasting spot prices has grown rapidly over the past few years. Examples include approaches in McCallum and Wu (2005) and Alquist and Kilian (2010). The evidence on whether using risk adjusted futures prices are useful in prediction is mixed. For instance, Moosa and Al-Loughani (1994) found evidence of a risk premium in crude oil futures markets and conclude that futures prices are not efficient forecasters of future spot prices. On the other hand, Pagano and Pisani (2009) found that the use of risk premium adjustment improves forecasting ability over long time horizons beyond six months. Apart from the models based on forecasting the futures prices, one can identify the following different types of oil spot price forecasting models in the literature.

- One can use spot price history directly to build a predictive model, e.g. based on a simple autoregressive moving average (ARMA) model structure. However, Reeve and Vigfusson (2011) show that futures-based forecasting typically outperforms such spot price based approaches.
- Another alternative for spot prediction is to use relevant macroeconomic variables other than spot price in a regression framework, leading to vector autoregression (VAR) models; see e.g. Baumeister and Kilian (2012) show that VAR models show good short term forecasting ability.
- Finally, structural models including a detailed structure of supply side dynamics are discussed in Kilian (2009), Kilian and Murphy (2012), Kilian and Murphy (2014), and Baumeister and Peersman (2013), among others.

Alquist et al. (2013) provide a comprehensive overview of different oil spot price forecasting methods.

There is a large literature on evaluating the forecasting performance of futures markets. Great research efforts have been expended in two areas: first is understanding the underlying mechanisms that determine the spot price and second is the development of many models suitable for forecasting the spot price. The idea of using oil futures prices to predict the spot price is based on the assumption that the futures prices reacts faster to the new information entering the market than the spot price. For example, Kumar (1992) presents evidence to support market efficiency and finds futures prices to be unbiased forecasters of crude oil prices. He investigates whether the forecasts from using futures prices can be improved by incorporating information from other forecasting techniques. Brenner and Kroner (1995) suggests that the inconsistencies observed between futures and spot prices may be as the result of carrying costs rather than a failing of the efficient market hypothesis. Girma and Paulson (1999) discovered that there are risk-arbitrage opportunities in petroleum futures spreads. Avsar and

Goss (2001) observe that inefficiencies are likely to be exacerbated in relatively young and shallow futures markets such as the electricity market, where forecast errors may indicate a market still coming to terms with the true market model. Moreover, Emery and Liu (2002) examine the relationship of futures prices between electricity and natural gas, and moreover, they report that there exist opportunities to profit from trading with futures spreads. According to Silvapulle and Moosa (1999) trading in the futures market has many advantages when compared to spot trading, such as low transaction cost, high liquidity, and low cash in up-front, among others. This makes it much more attractive for investors to react on new information than taking position in the spot market. There are several studies which have examined the modelling and forecasting of crude oil futures prices in recent years. For instance, Moshiri and Foroutan (2006) examined the chaos and non-linearity in crude oil futures prices. Performing several statistical and econometrical tests led them to conclude that futures prices time series is stochastic, and non-linear.

In addition, oil prices and its volatility both play vital roles in affecting the global economy. In the literature, several studies have found that higher oil prices have an unfavourable impact on the global economy (see for example, Morana (2013), Timilsina (2015), Archanskaïa et al. (2012)). In order to make appropriate decisions about the direction of economic policy, it is important to accurately forecast future oil prices with effective models (see, e.g. Hsu et al. (2016)). All the above approaches use price histories and/or numeric data on economic variables in the predictive model. In addition, there is information available in terms of global macroeconomic news which also has an impact on oil prices. The use of news or its proxy as an input to forecasting has been increasing in the last decade. Galati and Ho (2003) investigated to what extent daily movements in the euro/dollar rate were driven by news about the macroeconomic situation in the USA and the euro zone. Kim et al. (2004) investigated the impact of scheduled government announcements relating to six different macroeconomic variables on the risk and return of US bond, stock and foreign exchange markets. Arshanapalli et al. (2006) investigated the effects of macroeconomic news on time-varying volatility as well as time-varying covariance for the US stock and bond markets; they found that stocks and bonds have higher volatility on the day of macroeconomic announcements. Nikkinen et al. (2006) investigated how global stock markets are integrated with respect to the U.S. macroeconomic news announcements. Hess et al. (2008) investigated the impact of seventeen US macroeconomic announcements on two broad and representative commodity futures indices. Elder et al. (2012) examined the intensity, direction, and speed of impact of US macroeconomic news announcements on the return, volatility and trading volume of three important commodities: gold, silver and copper futures. Macroeconomic news is an important subset of all financial news; analysts qualitatively digest and apply this in financial decision making. Automatic analysis of macroeconomic announcements are now finding applications, in the asset classes of Fixed income, Foreign Exchange (FX) and Commodities. For example, Erlwein-Sayer (2017) has analysed the impact of macroeconomic news in predicting the yield spreads of the European sovereign bonds. Their preliminary findings confirm that this approach improves the analytic models for monitoring spreads.

In the present work, the Kalman filter has been utilised to build a model for forecasting arbitrage-free futures prices. Kalman filter was first proposed in Kalman (1960) as a recursive Bayesian estimator in engineering. Kalman filter is a method to reduce or eliminate the noise in experimental data. It has since been widely employed in applications outside engineering, especially in econometrics. For example, Schwartz (1997), Manoliu and Tompaidis (2002) and Lautier and Galli (2004) have applied Kalman filter to forecast spot prices with futures prices. A multi-commodity implementation is presented in Cortazar et al. (2008), where the futures prices of different commodities are used simultaneously to forecast the commodity prices. Inclusion of jumps to the commodity price process and a subsequent use of a particle filter for inference on commodity prices is advocated in Aiube et al. (2008). In Cortazar and Schwartz (2003), a three factor model for oil futures prices is suggested which departs from the Bayesian viewpoint used in filtering and infers prices using a numerical (but simple) optimisation instead. Mirantes et al. (2012) provide a model where the seasonality components are also stochastic and hence allow for a frequency variation. The Kalman filter and its various modifications have also been used within financial mathematics for modelling and forecasting of interest rates (Babbs and Nowman (1999), Date and Wang (2009)) and for estimating the asset price volatility from intra-day stock prices (Barndorff-Nielsen and Shephard (2002)). Linear filtering is used in Monoyios (2007) in the context of hedging in incomplete markets, for updating the estimates of uncertain drift parameters in the price process. Date and Ponomareva (2011) have provided a review of applications of filtering within finance.

In addition to the Kalman filter mentioned above, this study proposes to use quantified macroeconomic news sentiment to enhance the predictive power of the model. Crude oil prices are affected by many different macroeconomic events, including major industrial accidents, natural disasters, wars and political upheavals. This study uses news sentiment as an exogenous input which can change the volatility of the spot price. Specifically it uses the global macroeconomic news sentiment applied to a broad dataset of the crude oil prices. Empirical experiments have been carried out for forecasting the futures contract prices of crude oil using the global macroeconomic news sentiment. In Islyaev (2014)) a random parameter one factor model was proposed for commodity price modelling. We extend this approach by incorporating macroeconomic news sentiment. To the best of our knowledge, forecasting the futures prices of crude oil using global macroeconomic news sentiment has not been reported in the literature before. A linear state space model is used with logarithm of the spot price as the latent variable and a vector of logarithm of the futures prices as the observed variable, where this study looks at multiple futures with different term maturities. A one factor model is used with a constant risk premium, a random mean and a seasonality adjustment in terms of an additive sinusoid. This study uses logarithm of the prices of twelve futures contracts as observed variables. In addition, it assumes that the volatility of spot price depends on macroeconomic news (our model structure is described and justified later in section 5.3). The motivation for using a filtering-based framework stems from the fact that the futures market for commodities is more liquid than the spot market. Computational results of calibration and comparison for the models as well as the out-of-sample forecasting will be presented.

The rest of the chapter is organized as follows. Section 2 explains the two streams of time series, namely, the market data and macroeconomic news metadata that are used for numerical experiments. Section 3 describes the vector autoregressive (VAR) and the one factor models. Also a linear state space representation is defined for the one factor models. A new proposed model namely the macroeconomic news sentiment augmented model is presented in this section too. Section 4 describes the Kalman filter and the issues of model calibration, model fitting and the performance measures. Section 5 analyses the computational results of the empirical investigation of the 12 futures contracts of crude oil comparing the performances of the VAR model and the one factor model with and without the macroeconomic news data using the chosen performance measures. Section 6 sets out the conclusion.

5.2 Data

In this study, extensive numerical experiments have been conducted for calibration of different models and comparisons of their performance when it comes to forecasting crude oil spot and futures prices. Two different types of data has been used: market data about the prices of crude oil spot and crude oil futures, and global macroeconomic news data.

5.2.1 Market Data

The historical market data used to estimate and validate the parameters of the models (described later in section 5.3) consist of daily closing prices of WTI Crude Oil traded on New York Mercantile Exchange (NYMEX) from January 2, 2014 to December 21, 2016. The time series, which is obtained from Thomson Reuters, includes spot prices and 12 corresponding futures contracts which expire on various dates from June 20, 2017 to May 22, 2018. Table 5.1 presents some statistics for the daily spot prices of WTI Crude Oil. It is easy to observe that the spot prices are skewed to the right and are leptokurtic. These two observations may be explained by seasonal characteristics.

	Count	Mean	Std	Skewness	Excess Kurtosis	Min	25%	50%	75%	Max
Spot Prices	750	61.78	24.42	0.73	1.95	26.19	44.75	49.90	91.13	107.95

Table 5.1: Spot Price Statistics

Figure 5.1 shows both spot prices and futures prices of WTI Crude Oil. From January 2, 2014 to October 31, 2014, the futures prices are below spot prices while from November 3, 2014 to December 21, 2016, the futures prices are above spot prices. In futures markets, these situations are called as *backwardation* and *contango*, respectively. Normally, the spot prices are lower than the corresponding futures prices. Here, the Crude oil market experiences both backwardation and contango, which means that this market had experienced some unseen events during this time period.

Table 5.2 presents the corresponding relation between expiration dates and quotes for 12 futures contracts. For brevity, futures contracts including futures prices are represented with corresponding quotes.

Table 5.3 presents basic statistics of WTI Crude oil futures prices. From this table, one can observe that futures prices data are slightly positively skewed and their excess kurtosis are all about 1.75, which is very close to that for spot price data. This means that both spot price data and futures prices data have analogous statistical properties, and moreover there may be some seasonal patterns embedded in these data.

Currently, the futures contract with the farthest maturity date expires at December 2025. However, only the futures contracts with the latest 12 maturity dates are

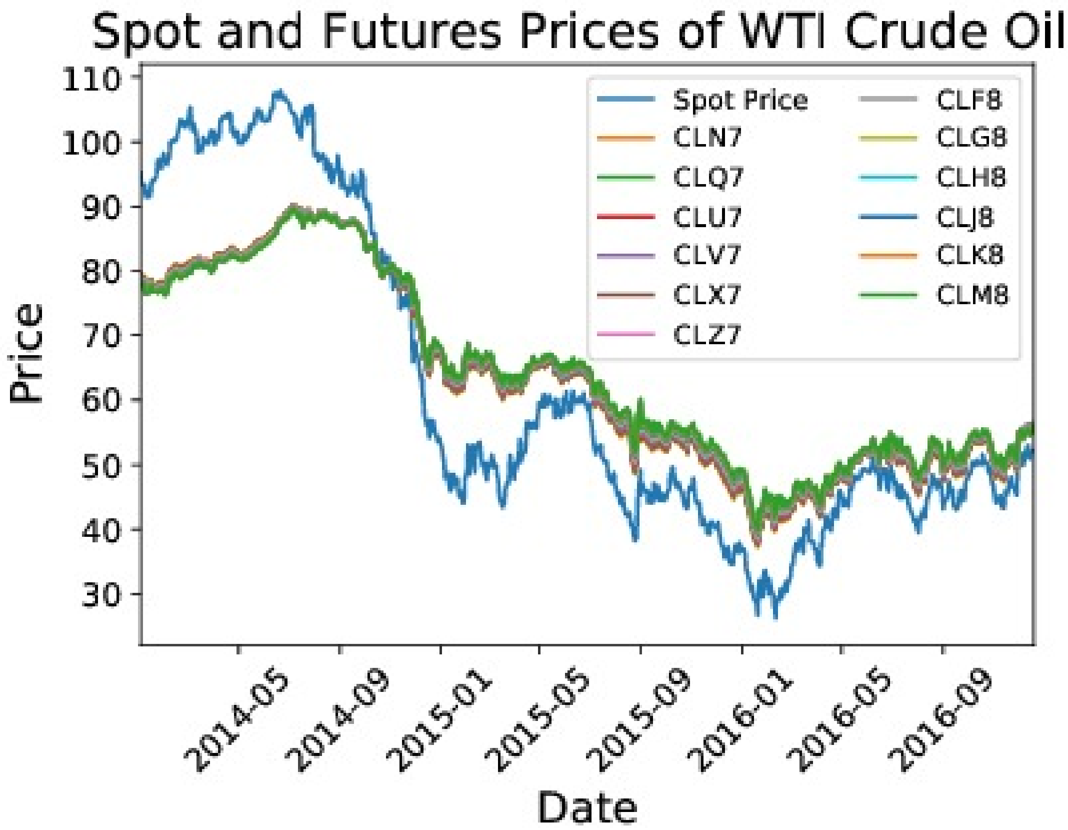

Figure 5.1: Spot and Futures Prices of WTI Crude oil

	Quote	Expiration Date		Quote	Expiration Date
1	CLN7	2017.06.20	**7**	CLF8	2017.12.19
2	CLQ7	2017.07.20	**8**	CLG8	2018.01.22
3	CLU7	2017.08.22	**9**	CLH8	2018.02.20
4	CLV7	2017.09.20	**10**	CLJ8	2018.03.20
5	CLX7	2017.10.20	**11**	CLK8	2018.04.20
6	CLZ7	2017.11.20	**12**	CLM8	2018.05.22

Table 5.2: Corresponding Relation between Expiration Dates and Quotes

Quote		**Count**	**Mean**	**Std**	**Skewness**	**Kurtosis**	**Min**	**25%**	**50%**	**75%**	**Max**
1	CLN7	750	63.11	15.24	0.39	1.74	37.22	50.51	60.82	79.78	90.27
2	CLQ7	750	63.2	15.11	0.39	1.74	37.47	50.69	61.06	79.79	90.16
3	CLU7	750	63.32	14.99	0.39	1.74	37.73	50.86	61.27	79.81	90.09
4	CLV7	750	63.46	14.89	0.38	1.74	37.97	51.05	61.45	79.75	90.06
5	CLX7	750	63.61	14.8	0.38	1.74	38.22	51.30	61.63	79.80	90.06
6	CLZ7	750	63.77	14.71	0.37	1.74	38.49	51.54	61.88	79.75	90.08
7	CLF8	750	63.81	14.61	0.37	1.74	38.66	51.66	61.96	79.66	89.98
8	CLG8	750	63.86	14.51	0.37	1.75	38.85	51.77	62.09	79.58	89.89
9	CLH8	750	63.92	14.41	0.37	1.75	39.06	51.91	62.27	79.47	89.81
10	CLJ8	750	64	14.31	0.37	1.75	39.30	52.04	62.45	79.41	89.73
11	CLK8	750	64.09	14.21	0.36	1.75	39.56	52.19	62.66	79.28	89.66
12	CLM8	750	64.19	14.11	0.36	1.75	39.83	52.36	62.87	79.21	89.60

Table 5.3: Futures Prices Statistics

considered in this study. Having all the futures maturing beyond the last date of spot price data set also avoids the problem of very high volatility of futures price when it nears maturity.

There are two types of dates in futures prices data. One is the trading date which refers to the calendar date. For example, the latest date of the historical futures price data is December 21, 2016, meaning that the data was collected by that day and there is no new data later than this date. The other type is contract date referring to the delivery date of the futures contract. e.g., the June 20, 2018 WTI Crude oil contract means that this futures contract is deliverable on June 20, 2018.

5.2.2 Macroeconomic News Analytics Metadata

5.2.2.1 Macroeconomic News Analytic Data

The RavenPack News Analytics delivers sentiment analysis and event data most likely to impact financial markets and trading around the world. The service includes analytics of more than 192,000 entities in over 130 countries and covers over 98% of the investable global market. All relevant news items about entities are classified and quantified according to their sentiment, relevance, topic, novelty, and market impact;

the result is a data product that can be segmented into many distinct benchmarks and used in a variety of applications. Since RavenPack macroeconomic data is the key to this research, their product is described in some details below.

In particular, RavenPack indicators can protect portfolio managers or traders from the consequences of missing important news that has an impact on their position or portfolio. They are built on top of RavenPack's core technology which continuously analyses relevant information from major real-time news-wires and trustworthy internet sources including more than 19,000 financial sites, blogs, and local and regional newspapers to produce real-time structured sentiment, relevance and novelty data for entities and events detected in the unstructured text.

RavenPack classifies news items using multiple sophisticated sentiment detection algorithms. In addition, RavenPack generates a number of non-sentiment analytics including information about companies, events, relevance and market impact. Outputs are often in the form of numerical news scores that can be used as inputs in the calculation of company, sector and industry indicators.

The RavenPack News Analytics data is divided into two parts, the equity (company) related analytics and the global macro analytics. Each record in the Global Macro News Analytics database contains 34 fields including a time-stamp, entity ID, entity name, entity type, topic, group, type, scores for relevance, novelty and sentiment, and unique identifiers for each news story analysed. In the historical metadata files, each row in the database represents an entity-level record. Thus, whenever an entity such as a company or currency is mentioned in the news, RavenPack produces an entity-level record. A single news story can yield multiple records if more than one entity is mentioned.

For each entity RavenPack assigns a unique ID. The entities are classified into different types, currently RavenPack supports the following 9 entity types:

1. COMP (Company): Business organization that may be traded directly on an exchange.

2. ORGA (Organization): Non-business organizations such as a government, central bank, not-for-profit, terrorist organization, etc.

3. CURR (Currency): Currencies of all financial and industrial countries.

4. CMDT (Commodity): Exchange traded commodities such as crude oil and soy.

5. PLCE (Place): Towns, cities and countries.

6. NATL (Nationality) The status of belonging to a particular nation.
7. PEOP (People) Individuals that are mentioned in the news.
8. PROD (Products) and services.
9. TEAM (Sports Teams) Professional teams from a variety of different sports.

Relevant stories about entities are classified into a set of predefined event categories following the RavenPack taxonomy. There are over 2,000 types of event categories automatically detected by RavenPack. They have been divided into different groups, such that each group contains a collection of related categories. Following is the RavenPack taxonomy fields:

- TOPIC: A subject or theme of events detected by RavenPack. The highest level of the RavenPack Event Taxonomy. There are 5 topics included in the 4.0 event taxonomy.
- GROUP: A collection of related events. The second highest level of the RavenPack Event Taxonomy. There are 51 groups included in the 4.0 event taxonomy.
- TYPE: A class of events, the constituents of which share similar characteristics. There are 412 types included in the 4.0 event taxonomy.
- SUB-TYPE: A subdivision of a particular class of events. There are 130 sub-types included in the 4.0 event taxonomy.
- PROPERTY: A named attribute of an event such as an entity, role, or string extracted from a matched event type. When applicable, the role played by the entity in the story is detected and tagged. There are 68 properties included in the 4.0 event taxonomy.
- CATEGORY: A unique tag to label, identify, and recognize a particular type and property of an entity-specific news event. There are 2,064 categories included in the 4.0 event taxonomy.

More details about entity types and taxonomy fields can be found in the RavenPack (2014) user guide. In the following subsection, we will briefly explain how the high volume of macroeconomic news data has been filtered in this research.

5.2.2.2 Choosing Macroeconomic News Items

This study has used macroeconomic news sentiment from a commercial data provider Ravenpack to enhance the predictive ability of the model, although the methodology is relevant for quantitative news data from other sources. First one has to decide which news item should be included into the macroeconomic news data. It is not practical to include all the available macroeconomic news items to generate positive and negative time series, which reflect the impact of macroeconomic news on the oil futures prices. One has to be careful in choosing the variables and try to eliminate the irrelevant news items from the macroeconomic data.

Most macroeconomic news items are released on different days and at different frequencies, making it difficult to process the flow of information in a systematic and consistent way. The actual news releases occur with a variety of different lags with respect to the month they are referencing. Furthermore, news on different events are frequently released simultaneously. Finally, the release frequency varies across different economic aggregates, for instance, data releases of different economic announcements are usually observed at different frequencies, e.g. quarterly, monthly, weekly, etc.

The aim in this study is to extract a set of factors that have a close relationship with crude oil price movements, rather than relying on all macroeconomic news flow that come from different entities over the world. A specific filter[1] is imposed to select the factors that are assumed to have impacts on the crude oil price movements and have an economically motivated structure on the macroeconomic news flow. The basic elements of how the high volume data from Ravenpack has been filtered is outlined below. The reader is referred to the RavenPack (2014) manual for further information on how the data is structured.

- Ravenpack data might be categorized according to different entity types. This study focuses on entities related to crude oil price movements: company, organization, currency, commodity, place, people and product. Macroeconomic news items coming only from these entity types is considered as a part of our input.
- For the chosen entity types, the data may also be categorized by chosen events. Out of 51 event categories available, the following have been chosen as they assumed to be related to crude oil price movements: “civil-unrest”, “commodity-prices”, “consumption”, “domestic-product”, “exploration”, “foreign-exchange”,

[1] The term *filtering* in this context is used in colloquial sense to denote appropriate selection of factors and has nothing to do with the Kalman filter described earlier.

"industrial-accidents", "interest-rates", "taxes", "transportation", "natural-disasters", "production", "products-services", "war-conflict".

- Finally, the fact that crude oil price fluctuations in the oil markets have a direct relationship with the news coming from the export and import countries, as supplier and demander of the crude oil, the macroeconomic data (for the above entity types and the above events) is filtered in such a way that only the top 15 export and import countries are taken into consideration in this research. Tables C.1 and C.2 in the *Appendix* C display the import and export countries of crude oil and their proportions in global contributions.

In a summary, the fairly simple data-selection method in this study delivers an approach to construct a daily time series that can be used as a proxy of global macroeconomic news data which can then be incorporated in the predictive models. Finally data is obtained on the dates, release times, sentiment and relevance scores, and more fields for a total of 27607 global entities. Macroeconomic news items covering the period from 2 January 2014 through 21 December 2016, for a total of more than 2,730,000 news items over about 750 business days are filtered. This data is obtained from RavenPack news analytics, which provides precisely time-stamped data. Table C.3 in the *Appendix* C shows the statistics summary of the news sentiments for each group.

After choosing macroeconomic news data in this fashion, Ravenpack news analytics generates time-stamped news sentiment scores, novelty scores and relevance scores for each news item, for the specified time period. The quantity of interest for this study is news sentiment scores, which range from +1 (most positive news) to -1 (most negative news), with 0 as a neutral score. The novelty score ranges from 0 (least novel) to 100 (most novel) and indicates the novelty of the news item, whereas the relevance score indicates relevance and also ranges from 0 (least relevant) to 100 (most relevant). This research restricts to news sentiment scores from those news items which have a novelty score of 80 or more and relevance score of 100. This study also separate out positive news sentiment scores and negative news sentiment scores, as two separate time series. Finally, this gives us the time-stamped news sentiment score. This needs to be processed further, as any day might still have multiple macroeconomic news items at different frequencies and crude oil price time series is on a day scale. The method to arrive at *macroeconomic news impact scores* from macroeconomic news sentiment scores is outlined in subsection 5.3.6.1.

5.3 Models for Spot and Futures Prices

This section presents a detailed description of dynamic models used in this work. The main emphasis is on the one factor model as discussed in Manoliu and Tompaidis (2002). Depending on whether there is a risk premium and a seasonality component or not, three different models have been used in this study, namely one factor model, one factor model with risk premium, and one factor model with risk premium and seasonality. These three models are described in the following subsections. Then a new method of incorporating macroeconomic news sentiment in a predictive model for forecasting prices of crude oil future contracts has been defined. Before moving on to these models, it is worth seeing a popular econometric model called vector autoregressive (VAR) model.

5.3.1 Vector Autoregressive Model

The VAR model is one of the most successful, flexible, and easy to use models for the analysis of multivariate time series. It is a natural extension of the univariate autoregressive model for dynamic multivariate time series. Sims (1980) first proposed the VAR approach that reduced the restrictions needed by earlier econometric models and allowed the data to be modelled in an unrestricted form, where all variables in the model are considered as endogenous. The VAR model has proven to be especially useful for describing the dynamic behaviour of econometric and financial time series and for forecasting. It often provides superior forecasts to those from univariate time series models and elaborates theory-based simultaneous equations models. Forecasts from VAR models are quite flexible because they can be made conditional on the potential future paths of specified variables in the model.

VAR models are used for multivariate time series. The structure is that each variable is a linear function of past lags of itself and past lags of the other variables. A VAR(p) model of the $m \times 1$ vector of time series $y_t = (y_{1t}, y_{2t}, \dots, y_{mt})$ with autoregressive order p is given by

$$y_t = Bx_t + \epsilon_t, \tag{5.1}$$

where

$$x_t = \begin{bmatrix} 1 & y_{t-1}^\top & \cdots & y_{t-p}^\top \end{bmatrix}^\top, \tag{5.2}$$

and the dimensions of B and x_t are $(m \times (mp+1))$ and $((mp+1) \times 1)$, respectively. ϵ_t is a $m \times 1$ vector of disturbances that have mean zero and covariance matrix Σ_ϵ.

The VAR model has its own advantages and disadvantages. The first advantage is that the model is easy to estimate even with commercial software. Another advantage is that the VAR model treats all variables in the system symmetrically as endogenous, so there is no need to specify which variables are endogenous or exogenous. Furthermore, the VAR model can assist identification of shocks, including monetary policy shocks. Finally, the VAR model has good forecasting capabilities. However, the model has also been criticized. The main criticism is that the model has a very large number of parameters. This leads to problems in estimation such as over-fitting and inconsistent results with different estimation schemes. When it comes to forecasting futures prices on the same underlying asset, a VAR model does not impose no-arbitrage condition on the forecasts, which is also seen as a limitation. This no arbitrage condition is automatically imposed if one considers all the futures prices driven by a single source of uncertainty affecting the spot price itself, as explained in the subsequent sections.

5.3.2 One Factor Model of Crude Oil Spot Price

To construct a one factor model for spot and futures prices of WTI Crude oil, the log of spot price S_t, $x_t = \log(S_t)$, is assumed to be driven by a stochastic process in the filtered probability space $(\Omega\ ,\ P,\ F_t)$. Here, Ω is the set that contains all the possible realisations of x_t, P is the objective probability measure and F_t is the natural filtration. For brevity, t_n, $(t_n : n = 0, \ldots, k, t_0 < t_1 < \cdots < t_k, \Delta := t_n - t_{n-1})$ is used to denote the discrete time.

Specifically, $x_t = \log(S_t)$ is assumed to follow a mean-reverting Ornstein-Uhlenbeck (OU) process:

$$dx_t = (\alpha - kx_t)dt + \sigma dW_t^P, \tag{5.3}$$

where κ is called the speed of mean reversion, α determines the long-run mean of x_t, and W_t is the Wiener process under objective probability measure P. Note that one factor model means that there is only one Wiener process in it. Further details about the Wiener process can be found in Øksendal (2003).

5.3.3 One Factor Model with Risk Premium

Oil futures prices include risk premia, like any other risky asset. This risk reflects the possibility that oil spot prices at the delivery time can be lower or higher than the contracted price. Therefore, the risk premium can be counted as a reward for holding a risky asset rather than a risk-free one. In a market free of arbitrage, a risk-neutral measure exists, and the process x_t satisfies the following stochastic differential equation(SDE) if it is generated by Wiener process under risk-neutral measure R:

$$dx_t = (\tilde{\alpha} - kx_t)dt + \sigma dW_t^R. \tag{5.4}$$

These two drift items are associated with $\tilde{\alpha} = \alpha - \lambda_t \sigma$ for some processes λ_t i.e., $dW_t^R = dW_t^P + \lambda_t dt$. The exact constraint on λ_t can be found in Duffie (2010). For the purpose of this study, λ_t is assumed to be constant. This is a very common assumption in the literature.

By applying Ito's lemma to the function $f(x_t, t) = e^{\kappa t} x_t$, it can be clearly shown that $x_{t+\Delta}$ has the following mean and variance conditional on a past value x_t, keeping in mind the process x_t is normally distributed under the risk-neutral measure:

$$E^R(x_{t+\Delta}|F_t) = x_t e^{-k\Delta} + \frac{\tilde{\alpha}}{k}(1 - e^{-k\Delta}), \tag{5.5}$$

$$Var^R(x_{t+\Delta}|F_t) = \frac{\sigma^2}{2k}(1 - e^{-2k\Delta}). \tag{5.6}$$

Here, the modification of the above model as suggested in Islyaev (2014) is followed to make it more flexible at very modest increase in complexity. Specifically, it is assumed $\alpha \sim N(\mu_0, \theta^2)$ in equation (5.3). This permits the logarithm of spot price to converge to a random mean and potentially improves the prediction ability of the model with only one parameter added based on the original model. The effect of adding this parameter in terms of improving the prediction ability of the model will be investigated later in the forecasting results section. This study also assumes the random mean to be uncorrelated with the Wiener process and hence has no risk premium. The expressions for conditional mean and variance can be re-expressed as:

$$E^R(x_{t+\Delta}|F_t) = x_t e^{-k\Delta} + \frac{\mu_0 - \lambda\sigma}{k}(1 - e^{-k\Delta}), \tag{5.7}$$

$$Var^R(x_{t+\Delta}|F_t) = \frac{\theta^2}{k^2}(1 - e^{-k\Delta})^2 + \frac{\sigma^2}{2k}(1 - e^{-2k\Delta}). \tag{5.8}$$

Now, define the futures contracts maturity dates as $T = (T_i : i = 1, \ldots, m, 0 < T_1 < T_2 < \cdots < T_m)$. Then the futures price of crude oil with log-spot price x_t at time $t < T_i$ for maturity T_i can be expressed as a conditional expectation of crude oil price at the maturity time of the futures contract: $F(t, T_i) = E^R(e^{x^i}|F_t), i = 0, \ldots, m$, where the expectation is taken under R measure and for brevity of notation, x_{T_i} is expressed as x^i. In addition, $F(t, T_i) > 0$ when $t > T_i$. The expiration time of the i^{th} futures contract is expressed as $\Delta_t^i = T_i - t$. Since S_t is assumed to be log-normally distributed, then

$$F(t, T_i) = E^R(e^{x^i}|F_t) = e^{E^R(x^i|F_t) + \frac{1}{2}Var^R(x^i|F_t)}. \tag{5.9}$$

Using the above equation one can derive an affine equation for the vector of log-futures prices in terms of the log-spot price as:

$$vec\{y_t^i\} = x_t e^{-k\Delta_t^i} + \frac{\mu_0 - \lambda\sigma}{k}(1 - e^{-k\Delta_t^i}) + \frac{\sigma^2}{4k}(1 - e^{-2k\Delta_t^i}) + \frac{\theta^2}{2k^2}(1 - e^{-k\Delta_t^i})^2, \tag{5.10}$$

where the vec operator is defined by $vec(z_i) = [z_1 \quad z_2 \quad \ldots \quad z_n]^T$ and $y_t^i = \log F(t, T_i)$. Note that, here the convenience yield is not modelled explicitly and is considered that it is already reflected in the prices of futures contracts. Again, this approach followed in Islyaev (2014) is consistent with the framework followed in Manoliu and Tompaidis (2002). In contrast, Hyndman (2007) modelled the convenience yield explicitly.

5.3.4 One Factor Model with Risk Premium and Seasonality

Since the futures prices of some commodities especially crop and energy commodities heavily depend on the weather conditions, it is worth taking some seasonal characteristics into consideration. In this section, a simple model for seasonality with a single sinusoid is considered, which is parametrized as:

$$f(t) = \exp(c_1 + c_2 \sin(c_3 t + c_4)), \tag{5.11}$$

where c_1, c_2, c_3 and c_4 are constants and representing constant level, the amplitude, the frequency, and the phase of a seasonal pattern, respectively. Accordingly, one can modify the prices of futures as follows:

$$F(t,T_i) = f(T_i)E^Q(e^{x^i}|F_t), \tag{5.12}$$

and

$$vec\{y_t^i\} = \log(f(T_i)) + x_t e^{-k\Delta_t^i} + \frac{\mu_0 - \lambda\sigma}{k}(1-e^{-k\Delta_t^i}) + \frac{\sigma^2}{4k}(1-e^{-2k\Delta_t^i}) + \frac{\theta^2}{2k^2}(1-e^{-k\Delta_t^i})^2, \tag{5.13}$$

which denotes a vector of log-futures prices, with i^{th} element of the vector denoting log-futures price for time to maturity Δ_t^i, as defined earlier. In practice, seasonality can be parametrised using multiple sinusoids. Nevertheless, this complicates parameter estimation without necessary improving the quality of out-of-sample price forecasting.

5.3.5 Linear State Space Representation for the Model

A state space representation is a mathematical model of a physical system as a set of input, output and state variables related by first-order differential or difference equations. In other words, it is the set of all possible states of a dynamical system, each state of the system corresponds to a unique point in the state space. State equations may be obtained from an n^{th} order differential equation or directly from the system, with each state corresponding to a unique point in the model by identifying appropriate state variables. One advantage of the state space method is that the form lends itself easily to the digital and analog computation methods of solution. Additionally, if the dynamical system is linear and time invariant, the differential and difference equations may be written in matrix form.

For those three models described in the previous subsections, a linear state space representation will be used with a measurement equation based on the observable time series of futures prices and a discretized transition equation of the logarithm of spot commodity price, which is assumed to be latent. This allows us to use the Kalman filter to estimate the parameters by constructing and maximising a likelihood function, and to predict the futures prices. The state space equations for one factor with risk premium and seasonality model in subsection 5.3.4 is provided below. The models without seasonality are obtained by setting the relevant parameters to zero.

The state space equations corresponding to the model in subsection 5.3.4 can be expressed as

$$x_{n+1} = Bx_n + g + Rw_{n+1} \tag{5.14}$$

$$y_n = A_n x_n + d_n + Qz_n \tag{5.15}$$

where the state space model parameters may be expressed in terms of the original model parameters as:

$$f(t_n) = c_1 + c_2 \ \sin(c_3 t_n + c_4) \tag{5.16}$$

$$B = e^{-\kappa\Delta} \tag{5.17}$$

$$g = \frac{\mu_0}{\kappa}(1 - e^{-\kappa\Delta}) \tag{5.18}$$

$$R^2 = \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa\Delta}) + \frac{\Theta^2}{\kappa^2}(1 - e^{-\kappa\Delta})^2 \tag{5.19}$$

$$A_n = \begin{pmatrix} e^{-\kappa\Delta_n^1} \\ \vdots \\ e^{-\kappa\Delta_n^m} \end{pmatrix} \tag{5.20}$$

$$d_n = \begin{pmatrix} \frac{\mu_0 - \lambda\sigma}{\kappa}(1 - e^{-\kappa\Delta_n^1}) + \frac{\sigma^2}{4\kappa}(1 - e^{-2\kappa\Delta_n^1}) + \frac{\Theta^2}{2\kappa^2}(1 - e^{-\kappa\Delta_n^1})^2 + f(T_1) \\ \vdots \\ \frac{\mu_0 - \lambda\sigma}{\kappa}(1 - e^{-\kappa\Delta_n^m}) + \frac{\sigma^2}{4\kappa}(1 - e^{-2\kappa\Delta_n^m}) + \frac{\Theta^2}{2\kappa^2}(1 - e^{-\kappa\Delta_n^m})^2 + f(T_m) \end{pmatrix}. \tag{5.21}$$

For brevity of notation, $\Delta_{t_n}^i$ is represented as Δ_n^i, and m is the number of futures prices available at each t_n. $Q = cI_m$, where c is a scalar constant indicating the standard deviation of measurements and I_m is an $m \times m$ identity matrix. μ_0, λ, σ, Θ and κ are as defined in subsection 5.3.2 and c_1, c_2, c_3 and c_4 are constants.

5.3.6 One Factor Models with Macroeconomic News Data

Before proposing a method of incorporating macroeconomic news sentiment in a predictive model for forecasting prices of crude oil future contracts, in the following subsection we illustrate the process of generating two daily time series (positive and negative) that reflect the impact of macroeconomic news on the oil futures prices.

5.3.6.1 Generating Macroeconomic Impact Scores

After extracting the macroeconomic news data that is related to the crude oil price movements in the first stage in subsection 5.2.2.2, one can select the macroeconomic news sentiment score for each news items based on their relevance and novelty scores. In this study, rather than using the macroeconomic news sentiment itself we are going to use the idea that was first proposed by Yu (2014) to construct news impact scores. We construct macroeconomic news impact scores that can be used as proxies of global macroeconomic news impact in the new model. To achieve this, the RavenPack's macroeconomic news analytics database is used and some of its quantitative sentiment scores are employed. In the second stage, the following steps have to be done:

1. The time-stamp for each news item has to be converted from UTC to EST time, which is the timing convention of crude oil futures data from the New York Mercantile Exchange (NYMEX).
2. Then data has to be filtered such that choosing the macroeconomic news items that have relevance score of 100, which indicates how strongly related the entity is to the underlying news story, and novelty score of 80 or greater.
3. Separating the positive and negative sentiment scores so that two different time series can be obtained.
4. After separating the scores, in a similar fashion to Yu et al. (2015), the positive and negative macroeconomic impact scores for each sentiment score are calculated.
5. Finally, two daily time series are generated that represent the positive and negative macroeconomic news impact scores.

In the last step one has to be careful in generating daily time series since it is possible to get more than a macroeconomic news item in a day. Thus, one has to find a way to produce a single positive score and a single negative score for each day. In this study, in order to generate two daily time series that represent the positive and negative macroeconomic news impact scores, in which each day ends up with only one positive score as well as one negative score, the following four different approaches have been attempted:

(i) Aggregating all positive and negative macroeconomic news impact scores separately within a day to generate a single daily score for each time series (positive and negative).

(ii) Taking the average of the aggregated scores, which was produced by the first approach, to produce a single score for each time series for that day.

(iii) Giving weights to each macroeconomic news impact score based on how new is the news item that the sentiment score was driven from.

(iv) The last approach is to give the highest weight to the last news item (here macroeconomic news impact score) in the day, in which a weight of 75% is given to the last macroeconomic news item and 25% to the rest of news items.

The following equations from (5.22) to (5.25) are representing the mathematical expressions for the aforementioned approaches, respectively.

$$\text{Aggregated} = \sum_{i=1}^{n} impact_i, \tag{5.22}$$

$$\text{Averaged} = \sum_{i=1}^{n} \frac{impact_i}{n}, \tag{5.23}$$

$$\text{Weighted} = \sum_{i=1}^{n} w_i * impact_i, \tag{5.24}$$

$$\text{Greater-Weight-To-Last-News} = 0.25 * \sum_{i=1}^{n-1} impact_i + 0.75 * impact_n, \tag{5.25}$$

where $\{impact_1, impact_2, \ldots, impact_n\}$ and $\{w_, w_2, \ldots, w_n\}$ are the macroeconomic news impact scores and their weights in the day, respectively.

Having generated the macroeconomic news impact scores (positive and negative) using the above approaches, the empirical results show that taking the averaged impact scores (equation 5.23) is the most appropriate approach to be used because it gives better out of sample performance in terms of the chosen error measures. Therefore, all the experiments reported in this study have used this approach to generate the two time series that represent the good and bad global macroeconomic news that impact the crude oil futures prices, and have been utilized as proxies of positive and negative macroeconomic news impact scores in the aforementioned models.

5.3.6.2 Macroeconomic News Sentiment Augmented Models

The aim of this study is to develop a forecast model for crude oil prices which enables us to use the relative information content of the global macroeconomic news data in order to improve the predictive power of the aforementioned models. In this study, rather than using the macroeconomic news sentiment itself the idea that was proposed in chapter 4 is used to construct macroeconomic impact scores. As derived in section 4.3.1, the macroeconomic impact sores are used as proxies of global macroeconomic news impact in the new models. To achieve this, the RavenPack's macroeconomic news analytics database is used and some of its quantitative sentiment scores are employed.

To model the impact of the global macroeconomic news on the crude oil spot and futures prices, we define a scaling factor of the volatility, where it is determined by utilising the macroeconomic news sentiment score in the following way. The intuition behind this particular model structure is the same as in the previous chapter (Chapter 4) and is presented below for the sake of completeness. The model structure needs to reflect the following economic realities: positive and negative macroeconomic news impact the volatility differently. Furthermore, positive news tends to reduce volatility whereas negative news tends to increase volatility. We also seek a functional specification which restricts the impact of news on volatility from below and above to realistic limits. Finally, we are seeking a parsimonious model with a small number of added parameters. Keeping these requirements in mind, we choose the following functional form (which was also employed in Chapter 4): First, a function of two variables x and y is defined:

$$f(x, y) = a + 0.5 * b \left(\frac{e^x - 1}{e^x + 1} - \frac{e^y - 1}{e^y + 1} \right), \tag{5.26}$$

where a and b are constants. This function $f(x, y)$ lies between $(a, a + b)$ for any non-negative values of x and non-positive values of y.

Second, two different time series $\{\mathcal{P}_t\}$ and $\{\mathcal{N}_t\}$ are created that represent the positive and negative macroeconomic impact scores, respectively. Finally, the following function is defined as a scaling factor of the volatility:

$$f(\mathcal{P}_t, \mathcal{N}_t) = a + 0.5 * b \left(\frac{e^{\rho \mathcal{P}_t} - 1}{e^{\rho \mathcal{P}_t} + 1} - \frac{e^{\gamma \mathcal{N}_t} - 1}{e^{\gamma \mathcal{N}_t} + 1} \right), \tag{5.27}$$

where 0.5 is a scaling factor of the function, and a, b, ρ and γ are parameters of the model. This is in keeping with the intuition behind our earlier work. For further details, see chapter 4.

The new model structure of the one factor models is chosen to be one with a direct multiplicative effect of macroeconomic news on the volatility, as it is more natural to consider changes in percentage will change the volatility of price relative to its "news-neutral" level, when looking at the macroeconomic news impact. The positive and negative news may amplify or attenuate volatility level. Therefore, each σ in the above models will be multiplied by the volatility scaling factor. In the model enhanced by macroeconomic news, equations (5.19) and (5.21) in section 5.3.5 are replaced by the following two equations:

$$R^2 = \frac{(\sigma * f(\mathcal{P}_t, \mathcal{N}_t))^2}{2\kappa}(1 - e^{-2\kappa\Delta}) + \frac{\Theta^2}{\kappa^2}(1 - e^{-\kappa\Delta})^2 \tag{5.28}$$

$$d_n = \begin{pmatrix} \frac{\mu_0 - \lambda(\sigma * f(\mathcal{P}_t, \mathcal{N}_t))}{\kappa}(1 - e^{-\kappa\Delta_n^1}) + \frac{(\sigma * f(\mathcal{P}_t, \mathcal{N}_t))^2}{4\kappa}(1 - e^{-2\kappa\Delta_n^1}) + \frac{\Theta^2}{2\kappa^2}(1 - e^{-\kappa\Delta_n^1})^2 + f(T_1) \\ \vdots \\ \frac{\mu_0 - \lambda(\sigma * f(\mathcal{P}_t, \mathcal{N}_t))}{\kappa}(1 - e^{-\kappa\Delta_n^m}) + \frac{(\sigma * f(\mathcal{P}_t, \mathcal{N}_t))^2}{4\kappa}(1 - e^{-2\kappa\Delta_n^m}) + \frac{\Theta^2}{2\kappa^2}(1 - e^{-\kappa\Delta_n^m})^2 + f(T_m) \end{pmatrix}. \tag{5.29}$$

In addition, a constraint on the upper bound of the interval range is needed that is specified by summing up the parameters ($a > 0$) and ($b > 0$) to keep the macroeconomic news impact related scaling in a reasonable range. As in the earlier work in the previous chapter, the choice used here is $0 < a + b \leq 2$, *i.e.* the impact of macroeconomic news is assumed to change the volatility of the prices at most by a factor of 2.

The chosen model structure adds only four more model parameters and offers a reasonable compromise between increased model complexity and parsimony in terms of model parameters. The choice of model structure is essentially heuristic, and is justified through numerical experiments as well as prior experience (see Chapter 4).

One can also reduce the number of parameters by keeping the value of a and b fixed, for instance, $a = 0.5$ and $b = 1.5$. However, treating a and b as free parameters does improve results in terms of predictive power of the models.

5.3.7 Summary Forms for Computational Models

The main part of this study, the performance of two sets of computational models (the one factor model and the one factor model with macroeconomic news sentiment data) have been compared. These two sets of the models are explained in the earlier sections and summarized below:

- **One factor models:**
  The model is given by equations (5.14)-(5.21).
- **One factor models with macroeconomic news sentiment data:**
  The model is given by equations (5.14)-(5.18), (5.20), (5.28) and (5.29).

In both sets of models, futures prices are treated as observable variables and the spot price is treated as a latent variable. The methodology to estimate the parameters of the models in both cases is described in the next section. This study has also compared the one factor models with the VAR model described in subsection 5.3.1, and the results for this comparison are reported in the subsection (5.5.1). The VAR model is given by equation (5.1), and its endogenous variables are crude oil spot price and twelve corresponding futures prices.

5.4 Methodology

In this study, the linear Kalman filter and the maximum likelihood estimation method are employed to estimate parameters of the three aforementioned models. The following subsections give a more detailed explanation for these two techniques.

5.4.1 Kalman Filter

Kalman (1960) published his famous paper describing a recursive solution to the discrete data linear filtering problem. Its findings have become a solid base of extensive research and found widespread usage in applications in different areas as diverse as tracking, navigation, GPS, econometrics, autonomous, engineering and finance. As pointed out in Anderson and Moore (1979), the great success of Kalman filter is due to its modest computational requirement, elegant recursive properties, and its status as the optimal estimator for one-dimensional linear systems with Gaussian error statistics. The Kalman filter is a set of mathematical equations that can be used to estimate the state of the latent state of a process recursively using noisy observations, in a way that minimizes the trace of the covariance matrix of the estimation error. The filter can estimate the present and future states and it is possible to get the estimation even when the accurate information of the modelled system is unknown.

Filtering theory concerns the estimation of a latent state of a physical process from observations which may be corrupted by noise. In case of a discrete-time process, where the process is assumed to evolve in discrete steps of time, the mathematical model is often a random difference equation. For the discrete time case, the state is denoted as x_n and the observations or measurements as y_n. If it is desired to predict x_{n+j}, $j \in \mathbb{Z}^+$ (positive integers), based on observations up to time n, the problem is called prediction problem. If it is desired to obtain a probabilistic estimate of x_n based on observations up to time n, it is called a filtering problem. If it is desired to obtain a probabilistic estimate of x_{n-j}, $0 < j < n$, based on observations up to time n, the problem is called smoothing problem. If the random equation is linear in terms of the state of the process, the model is referred to as a linear model. The Kalman filter model, which is a discrete-time linear model and henceforth referred to as the "state equation", assumes the true state at time $(n + 1)$ is evolved from the state at n according to

$$x_{n+1} = Bx_n + g + Rw_{n+1}, \tag{5.30}$$

where B is the state transition model which is a linear function of time and applied to the previous state x_n, g is a vector and can also be a function of time, w_{n+1} is the process noise or is called the state disturbance error and R is the variance matrix and can also be a function of time.

The initial state x_0 and the state disturbance error w_{n+1} are assumed to be random vectors. At time n an observation (or measurement) y_n of the true state x_n is made according to

$$y_n = A_n x_n + d_n + Qz_n \tag{5.31}$$

where A_n and $Q_n \geq 0$ are known matrix-valued functions of time, d_n is a known vector valued function of time and $\{z_n\}$ is a sequence of i.i.d. variables with zero mean and identity matrix as covariance.

For the latent state equation (5.30) and the measurement equation (5.31), Kalman filter equations for estimating the mean and the covariance matrix of the latent variable can be summarised as follows:

Predict phase:

Predicted state estimate $$\hat{x}_{n+1|n} = B\hat{x}_{n|n-1} + g, \tag{5.32}$$

Predicted variance estimate $$P_{n+1|n} = BP_{n|n-1}B^T + RR^T - BP_{n|n-1}A_n^T S_n^{-1} A_n P_{n|n-1} B^T \tag{5.33}$$

Update phase:

Innovation or measurement residual $$V_n = y_n - (A_n\hat{x}_{n|n-1} + d_n) \tag{5.34}$$

Innovation (or residual) variance $$S_n = A_n P_{n|n-1} A_n^T + QQ^T \tag{5.35}$$

Optimal Kalman gain $$K_n = P_{n|n-1} A_n^T S_n^{-1} \tag{5.36}$$

Updated state estimate $$\hat{x}_{n|n} = \hat{x}_{n|n-1} + K_n V_n \tag{5.37}$$

Updated estimate variance $$P_{n|n} = (I - K_n A_n) P_{n|n-1} \tag{5.38}$$

$\hat{x}_{n|n-1}$ is the best estimated value of x_n at time n given observations up to time $\{n-1\}$ and $P_{n|n-1}$ is the corresponding conditional variance of $\hat{x}_{n|n-1}$ (a measure of the estimated accuracy of the state estimate). It is assumed that $\hat{x}_{0|-1}$ and $P_{0|1}$ are known.

In the present context, Kalman filter allows us to predict the spot price x_n as a hidden state, which also gives an arbitrage-free prediction of the vector of futures prices (as $A_n\hat{x}_{n|n-1} + d_n$).

Equations (5.32) and (5.33) provide an estimate of the mean and the covariance matrix of x_{n+1}, given information up to and including time t_n. Note that if sampling time is 0, $x_{0|0}$ can take any value, while $P_{0|0}$ can be a zero matrix if the initial value for a state is known or it can be assumed to be any diagonal matrix if the initial value is unknown. R is the process covariance or the prediction covariance. Put differently, this is the variance that will occur during the prediction stage. From equations (5.35) and (5.36), these are just the calculation of the Kalman gain from various covariances.

This is also the calculation based on the degree of trust, which is later used to update the predicted state variables. Q is the measurement covariance where measurements do not always give correct data. Obtaining the measurement covariance is also quite impractical, unless there is another accurate measurement to get the variance of each sensor. Sometimes the tuning of Q will depend on intuition. Equation (5.37) is actually the correction done to the state variable from the predicted state variable by adding the product of the Kalman gain and the residual. Equation (5.38) is the update of the state variable covariance.

5.4.2 Maximum Likelihood Estimation

5.4.2.1 Maximum Likelihood Estimation For One Factor Models

For the given log-futures prices measurements $F = \{y_1, \ y_2, \ \ldots, \ y_N\}$ up to time t_N, the maximum likelihood estimation algorithm (as described in subsection 2.2.2) can be applied along with the Kalman filter to calibrate parameters of equations from(5.16) to (5.29). Now the joint likelihood function for F can be written as

$$\hat{L}(F) = p(y_1) \prod_{i=2}^{N} p(y_i | F_{i-1}). \tag{5.39}$$

Since the measurements are jointly normal, the logarithm of the likelihood can be written as:

$$\log \hat{L}(F) = -\sum_{i=1}^{N} (\log |S_i| + V_i^T S_i^{-1} V_i), \tag{5.40}$$

where V_n and S_n are as defined in (5.34) and (5.35), respectively. The constant terms are ignored since they have no influence on the model when estimating the model parameters. For a given vector-value time series $\{y_1, \ y_2, \ \ldots, \ y_N\}$ and a vector of unknown model parameters θ, the optimisation problem can be stated as following:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}} (\log \hat{L}(F)). \tag{5.41}$$

Then $\hat{\theta}$ is used for forecasting experiments.

We implement the maximisation using Matlab's in-built routine *fminsearch*, which relies on the Nelder-Mead method.

5.4.2.2 Maximum Likelihood Estimation For VAR Model

To derive the maximum likelihood estimation of the VAR model, let us first consider a VAR(p) model of the $m \times 1$ vector of time series $y_t = (y_{1t}, y_{2t}, \ldots, y_{mt})$ with autoregressive order p:

$$y_t = c + A_1 y_{t-1} + \ldots A_p y_{t-p} + \epsilon_t, \tag{5.42}$$

where A_i are $m \times m$ coefficients matrices and c is a $m \times 1$ of intercepts. ϵ_t is a $m \times 1$ vector of disturbances that have mean zero and covariance matrix Σ_ϵ. VAR(p) can be written as:

$$y_t = Bx_t + \epsilon_t, \tag{5.43}$$

$$\text{where} \quad B = (c, A_1, \ldots, A_p) \tag{5.44}$$

$$\text{and} \quad x_t = \begin{bmatrix} 1 & y_{t-1}^\top & \cdots & y_{t-p}^\top \end{bmatrix}^\top . \tag{5.45}$$

The dimensions of B and x_t are $(m \times (mp + 1))$ and $((mp + 1) \times 1)$, respectively. Because the disturbances are normal distributed, the conditional density is multivariate normally distributed:

$$y_t | y_{t-1}, \ldots, y_{t-p} \sim N(Bx_t, \Sigma_\epsilon)'$$

and the conditional density of the t^{th} observation is:

$$= (2\pi)^{-(\frac{m}{2})} |\Sigma_\epsilon^{-1}|^{\frac{1}{2}} exp[(-\frac{1}{2})(y_t - Bx_t)' \Sigma_\epsilon^{-1} (y_t - Bx_t)]. \tag{5.46}$$

The likelihood function is the product of each one of these densities for $t = 1, \ldots, T$. The log-likelihood function is the sum of the log of all these densities. Since the disturbances are assumed to be normally distributed, the log likelihood in this case is:

$$l(B, \Sigma_\epsilon) = -(\frac{Tm}{2}) \log(2\pi) + (\frac{T}{2}) \log |\Sigma_\epsilon^{-1}| - (\frac{1}{2}) \sum_{t=1}^{T} \left[(y_t - Bx_t)' \Sigma_\epsilon^{-1} (y_t - Bx_t) \right] . \tag{5.47}$$

The value of B that maximise the log-likelihood happens to be the same as the

OLS estimator:

$$\hat{B} = \left[\sum_{t=1}^{T} y_t x_t^{'}\right] \left[\sum_{t=1}^{T} x_t x_t^{'}\right]^{-1}. \tag{5.48}$$

This means that the ML estimator of the VAR coefficients is equivalent to the OLS estimator of y_{jt} on x_t , that is, it is equivalent to applying OLS for each equation of the VAR separately. The OLS estimator for each separate equation is also equivalent to the system (multivariate) estimator. The ML estimator for the variance is:

$$\hat{\Sigma}_{\epsilon} = \left(\frac{1}{T}\right) \Sigma \hat{\epsilon}_t \hat{\epsilon}_t{}' \tag{5.49}$$

where

$$\hat{\epsilon}_t \equiv y_t - \hat{B} x_t. \tag{5.50}$$

5.4.3 Statistical Performance Measurements

In this study, we use the same measures for out-of-samples comparison as in the previous chapters, although these are repeated here for the sake of completeness. Out-of-sample comparison is carried out using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) which are defined respectively as:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |f_i - y_i| = \frac{1}{n} \sum_{i=1}^{n} |e_i| \tag{5.51}$$

and

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (f_i - y_i)^2}, \tag{5.52}$$

where f_i is the predicted value and y_i is the observed value. Since the errors are squared before they are averaged, the RMSE gives a relatively high weight to large errors.

On the other hand, for the in-sample comparison one might ask how to define the best model if there are more than one candidate model available? In such case, the model selection methods help us to select the best model. Therefore, given a dataset, the objective is to define which of the available models can best approximates the data. This involves trying to minimize the loss of information. The Akaike Information Criterion (AIC) has been widely used in selecting model for in-sample comparison of the increase in explanatory power proportional to increase in the model complexity.

The AIC, as described earlier in section 4.4.1, is an approach of selecting the best model, in terms of fit to the data, among other models. It is defined as:

$$\mathrm{AIC}(\hat{\Theta}) = -2(\log(\textit{likelihood})) + 2M \tag{5.53}$$

where $\hat{\Theta}$ is the set (vector) of model parameters, *likelihood* is the probability of the data given the candidate model and M is the number of estimated parameters in the candidate model (length of $\hat{\Theta}$).

An alternative method for in-sample comparison would be the second order information criterion (AICc), which takes into account the sample size by increasing the relative penalty for the model complexity with small datasets. It is defined as:

$$\mathrm{AICc}(\hat{\Theta}) = -2(\log(\textit{likelihood})) + 2M + \frac{2M(M+1)}{N-M-1} \tag{5.54}$$

where N is the sample size. As the number of data points (N) gets larger, the value of AICc converges to the value of AIC and this is because the denominator of the last term ($N - M - 1$) in the above equation (5.54) tends to approach N when the number of data points (N) becomes much greater than the number of parameters (M), and therefore, $2M(M + 1)/(N - M - 1)$ approaches zero. When the number of observations is quite small compared to the number of parameters, such that N/M is less than 40, the use of the second-order corrected AIC (AICc) is recommended by Burnham and Anderson (1998). There's really no disadvantage in always employing AICc regardless of the number of data points.

5.5 Computational Results

5.5.1 Comparison between VAR and One factor models

In this study, the daily closing prices of twelve crude oil futures contracts have been used. For each future's contract, 500 data points are used as in-sample data, which covers data from 2 January 2014 to 31 December 2015, to estimate the models parameters. 250 data points are used as out-of-sample data, which covers data from 2 January 2016 to 31 December 2016, to asses the forecasting performance of the models. From figure 5.1 one can see that during the past few years the crude oil market has experienced both backwardation and contango, both of which are totally different, meaning that this market has gone through unpredictable events.

First, a thirteen-dimensional VAR model for the endogenous variables (crude oil spot price and twelve corresponding futures prices) has been utilized as an alternative forecasting model to the one factor models. Having specified the model, the appropriate lag length of the VAR model has to be decided. In deciding the number of lags, the Akaike information criteria has been used, which is a common statistical method to use. Having specified the thirteen variables and the appropriate lag length of three, the VAR model has been estimated using the maximum likelihood estimation method under the assumption of normality of the errors. For the one factor models of the form (5.14)-(5.15), the Kalman filter and maximum likelihood are employed to estimate the models parameters. The same data has been used for the one factor models that was used in the VAR model, such as using the crude oil spot price as a latent variable in the state equation (5.14) and a vector of its twelve corresponding futures prices as observed variable in the measurement equation (5.15) at each time step. After calibrating the four models, they have been compared in terms of their forecastability ability using the statistical measurement mentioned in subsection 5.4.3. Table 5.4 shows the average errors of the predicted values of the twelve futures prices. The results indicate that the one factor models have outperformed the VAR model by a factor of 5 or more, on both the error criteria. Therefore, the decision to utilise the one factor models to construct new predictive models, which incorporate the global macroeconomic news data in their structures, is made.

	Models	**MAE**	**RMSE**
1	VAR	26.18	27.22
2	One Factor	3.74	4.57
3	One Factor with Risk premium	4.21	5.03
4	One Factor with Risk premium and Seasonality	2.24	2.86

Table 5.4: Out of sample comparison of four forecasting models in terms of MAE and RMSE errors.

Table 5.5 shows the complexity for the three models. Here, model complexity is simply defined as the number of model parameters. The complexity for one factor model is six. These six model parameters are: μ_0, Θ, κ, σ, Q and P_0, where both μ_0 and Θ are parameters of a normal distribution, the mean-reverting level is given by equations (5.3) and 5.4, α determines the long-run mean of the process, κ is the speed of reversion, σ is the volatility of the corresponding mean-reverting process, Q is a constant scalar which decides the standard deviation of the measurements, and

P_0 is the initial conditional variance. For one factor model with risk premium, its model complexity is seven. Besides the aforementioned six parameters, λ is added in the measure change. The complexity for one factor model with risk premium and seasonality is 11. On the basis of one factor model with risk premium, another four constant parameters which are c_1, c_2, c_3 and c_4 are added to simulate a seasonal function. They represent the constant level, the amplitude, the frequency and the phase separately.

	Models	**No. of parameters**
1	One Factor	6
2	One Factor with Risk premium	7
3	One Factor with Risk premium and Seasonality	11

Table 5.5: Complexity of the Models

5.5.2 Estimation Results

In this section, the empirical investigation of the estimated models for the 12 contracts of oil futures prices is presented. The in-sample results that are obtained by the one factor models which are described in section 4.3 are analysed and compared. As mentioned earlier in subsection (5.2.2.1), RavenPack's news analytics database has been used and some quantitative scores of its fields are utilised in this research.

After calculating the macroeconomic news impact scores for each sentiment score and generating two daily time series that represent the positive and negative macroeconomic news impact scores, the three models (the one factor, one factor with risk premium, and one factor with risk premium plus seasonality models) have been fitted to the observed time series of 12 futures prices via the Kalman filter and maximum likelihood method under the assumption of Gaussian error distribution. Thereafter the macroeconomic news impact scores are incorporated to the models and recalibrated to re-estimate the model parameters including those which come from the volatility scaling factor in equation (5.27).

In order to estimate the model parameters (Θ), the following steps have been carried out. First an initial value (guess) was given to each parameter in Θ. Then the Kalman filter recursions have been run over the whole time interval of the observable time series data for a specified value of the model parameters until an appropriate convergence criterion has been satisfied. Thus, after each recursion the log-likelihood function has been evaluated and an optimization of the procedure assigned to this particular problem and new estimates for the model parameters (Θ) are obtained through this. At each iteration the value of the log-likelihood function was compared against the value in the previous iterations and, if the difference between the current value and the old value was positive and smaller than a specified tolerance value (say 1e-12) the iterative procedure was stopped. Table C.4 in the *Appendix* C shows the parameter estimates of the models (with and without the macroeconomic data) for the 12 futures contracts of crude oil. The *fminsearch* function in MATLAB software is used to estimate the parameters. For the sake of brevity and clarity, simple names are given to the models that have been estimated in this study:

- Model 1: One factor model.
- Model 1-M: One factor model using Macroeconomic data.
- Model 2: One factor model with Risk premium.
- Model 2-M: One factor model with Risk premium using Macroeconomic data.
- Model 3: One factor model with Risk premium plus Seasonality.
- Model 3-M: One factor model with Risk premium plus Seasonality using Macroeconomic data.

To compare the six models in terms of goodness of fit to the data, first the Log-Likelihood values of the models are calculated and then the second order Akaike information criterion (AICc) is applied. The value of the AICc alone for a given time series is meaningless. Nevertheless, it becomes useful when it is compared to other AICc values of some other models. The model with the lowest AICc is described as the "best" model among all other models specified for the dataset. After applying this powerful tool in comparing various one factor models, the results show that the models with macroeconomic news sentiment seem to result a better fit to the data. Table 5.6 presents the values of the log-likelihood and AICc for the estimated models. According to the AICc values of the six models, an AICc ranking of the estimated models suggest the following ranking from the best to the worst: Model 3-M, Model 3, Model 1-M, Model 2-M, Model 1 and then Model 2.

Models	Log-Likelihood	AICc
Model 1	54848	-1.10E+05
Model 1-M	50171	-1.18E+05
Model 2	59767	-1.00E+05
Model 2-M	59107	-1.14E+05
Model 3	57173	-1.20E+05
Model 3-M	61458	-1.22E+05

Table 5.6: The log-Likelihood and AICc value of the estimated models

5.5.3 Forecasting Results

This section presents the results of forecasting crude oil futures prices for twelve future contracts using the one factor model with and without macroeconomic news data[2]. First the prior information of global macroeconomic news data is collected based on some filtering in the data to take the noise out. Then two time series have been created that represent the positive and negative impact scores. To estimate the models the Kalman filter and the maximum likelihood estimation method are employed. The forecasting method in this study is dynamic, where the predicted future prices are used for lagged future prices instead of the real prices when forecasting the next period.

In this study, the futures contracts whose maturities are between 1-month and 12-months are specifically chosen. This is because the research can cover short term, mid-term and long term futures prices. The prices of long term contracts have a much stronger dependence on macroeconomic news sentiment, as compared to futures with short term maturities. Incorporating a broad span maturities allows us to capture the market sentiment about the price evolution across the corresponding time spans.

As an alternative to the one factor models, the vector autoregressive (VAR) model was proposed and estimated in this study, in order to see the predictive power of the models. As demonstrated earlier in section 5.5, the VAR estimation results

[2] In this study, forecasting the crude oil spot price using the one factor model with and without macroeconomic news has been investigated as well but has not been reported. It was found that even though the one factor model with macroeconomic news performed better than the one factor model without macroeconomic news in terms of MAE and RMSE for the spot price but it was not significantly better.

show less forecasting accuracy than the one factor models. Thus, all experiments were carried out using the one factor models only. See Moon (1997) for further drawbacks of VAR based forecasting.

To asses the predictive power of the one factor models before and after incorporating the global macroeconomic news data, the out-of-sample performance of the one factor models with their macroeconomic versions are compared using some statistical measurements. The out-of-sample forecasting accuracy is mainly evaluated by two measures MAE and RMSE, which were defined earlier in subsection 5.4.3. The benchmarks in this study are the futures prices of crude oil. The MAE and the RMSE for each of the candidate one factor models (with and without incorporating macroeconomic news data) are presented in tables 5.7, 5.8 and 5.9. From these tables, one can clearly see that in all cases the models with macroeconomic news data appear to outperform and give better prediction than the models which had not incorporated the macroeconomic news data in their structures.

The first notable result from the tables is that the macroeconomic versions of the one factor models have performed better for the long-term futures contracts than the short- and mid-term contracts, although an overall improvement in predicting the prices of all the futures contracts is noticeable. The second notable result is that the macroeconomic versions of the one factor and the one factor model with risk premium plus seasonality appear to outperform the one factor model with risk premium, particularly for the long-term future contracts. Thus, the tables incidentally show the importance of modelling seasonality, which yields lower errors than models without seasonality.

The final concern is to see how much information could be extracted from each table about the range and average of the errors for the future contracts. For the one factor model (table 5.7) the MAE is ranged between 3.67 and 3.82, and the RMSE is ranged between 4.63 and 4.44, whereas the MAE and RMSE of its macroeconomic version are ranged between 1.7 and 2.12, and between 2.2 and 2.71, respectively. For the one factor model with risk premium (table 5.8) the MAE is ranged between 4.05 and 4.27, and the RMSE is ranged between 4.78 and 5.17, whereas the MAE and RMSE of its macroeconomic version are ranged between 1.69 and 2.32, and between 2.17 and 2.89, respectively. For the one factor model with risk premium plus seasonality (table 5.9) the MAE is ranged between 1.98 and 2.53, and the RMSE is ranged between 2.57 and 3.15, whereas the MAE and RMSE of its macroeconomic version are ranged between 1.8 and 2.07, and between 2.33 and 2.67, respectively. Both MAE and RMSE for each of the 12 contracts is better for the news augmented model as compared to the model which does not use news sentiment. Subsequently, the averages of the errors for the models are calculated. It is observed that the MAE and RMSE for the one factor model averaged across all the 12 future contracts are 3.74 and 4.57, respectively. Whereas the values MAE and MRSE are lower after incorporating the macroeconomic news data to the model, as they are 1.87 and 2.42, respectively. For the second model, the one factor model with risk premium, the MAE and RMSE averaged across all the future contracts are 4.21 and 5.03, respectively, whereas these values are drastically decreased to the half after incorporating the macroeconomic news data to the model, as 2 and 2.52 for MAE and RMSE, respectively. Finally, the values of MAE and RMSE of the one factor model with risk premium plus seasonality averaged across all the future contracts are 2.24 and 2.86, respectively. Once again, these error values are significantly lower after incorporating the macroeconomic news data to the model which results in 1.88 for MAE and 2.46 for RMSE. Table 5.10 summaries the above explanation.

Futures	**MAE**		**RMSE**	
	without	**with**	**without**	**with**
Futures 1	3.70	**2.12**	4.44	**2.71**
Futures 2	3.71	**2.07**	4.47	**2.64**
Futures 3	3.74	**2.03**	4.52	**2.59**
Futures 4	3.78	**2.00**	4.57	**2.55**
Futures 5	3.82	**1.96**	4.62	**2.50**
Futures 6	3.79	**1.89**	4.61	**2.43**
Futures 7	3.75	**1.81**	4.59	**2.34**
Futures 8	3.73	**1.77**	4.59	**2.30**
Futures 9	3.73	**1.74**	4.60	**2.27**
Futures 10	3.71	**1.71**	4.62	**2.24**
Futures 11	3.71	**1.70**	4.63	**2.22**
Futures 12	3.67	**1.70**	4.62	**2.20**

Table 5.7: MAE and RMSE for One Factor Model, with and without macroeconomic news data. The lower of two errors is highlighted in boldface.

Futures	**MAE**		**RMSE**	
	without	**with**	**without**	**with**
Futures 1	4.05	**2.32**	4.78	**2.89**
Futures 2	4.09	**2.25**	4.84	**2.82**
Futures 3	4.14	**2.21**	4.91	**2.77**
Futures 4	4.20	**2.16**	4.98	**2.72**
Futures 5	4.27	**2.12**	5.05	**2.67**
Futures 6	4.26	**2.05**	5.06	**2.57**
Futures 7	4.23	**1.96**	5.06	**2.46**
Futures 8	4.23	**1.90**	5.08	**2.39**
Futures 9	4.24	**1.84**	5.11	**2.34**
Futures 10	4.25	**1.79**	5.14	**2.28**
Futures 11	4.26	**1.74**	5.17	**2.23**
Futures 12	4.24	**1.69**	5.17	**2.17**

Table 5.8: MAE and RMSE for One Factor Model with risk premium, with and without macroeconomic news data. The lower of two errors is highlighted in boldface.

Futures	**MAE**		**RMSE**	
	without	**with**	**without**	**with**
Futures 1	2.53	**2.07**	3.15	**2.67**
Futures 2	2.46	**2.02**	3.09	**2.61**
Futures 3	2.42	**1.99**	3.06	**2.58**
Futures 4	2.39	**1.96**	3.02	**2.55**
Futures 5	2.35	**1.93**	2.99	**2.52**
Futures 6	2.29	**1.88**	2.91	**2.46**
Futures 7	2.20	**1.82**	2.81	**2.40**
Futures 8	2.14	**1.79**	2.75	**2.37**
Futures 9	2.10	**1.78**	2.71	**2.35**
Futures 10	2.06	**1.78**	2.67	**2.34**
Futures 11	2.02	**1.79**	2.63	**2.34**
Futures 12	1.98	**1.80**	2.57	**2.33**

Table 5.9: MAE and RMSE for One Factor Model with risk premium and Seasonality, with and without macroeconomic news data. The lower of two errors is highlighted in boldface.

	Models	**MAE**		**RMSE**	
		without	**with**	**without**	**with**
1	One factor model	3.74	1.87	4.57	2.42
2	One factor model with Risk premium	4.21	2.00	5.03	2.52
3	One factor model with Risk premium plus Seasonality	2.24	1.88	2.86	2.46

Table 5.10: Averages of the measurements errors across all the 12 future contracts for the one factor models with and without macroeconomic data

5.6 Summary

The first and the most important contribution of this study is that we demonstrate how to process macroeconomic news sentiment and how to incorporate it systematically into the commodity price model, in order to improve the forecasting of prices of crude oil futures. Our numerical experiments provide a strong evidence that macroeconomic news sentiment adds value to futures price forecasting. While most existing numerical studies only focus on short term futures contracts, our study includes the use of both short and long term futures contracts in model calibration as well as price forecasting.

The results of our study has widespread uses for public and private sector entities which rely on crude oil, and also for economic policy makers who need to incorporate future oil price behaviour into their economic scenarios. The proposed methodology is clearly applicable to other commodities where futures prices are more liquid that the underlying commodity price, and where macroeconomic news data is available.

Chapter 6

Conclusions and Future Research

This chapter concludes the thesis by summarizing the main contributions and outlines directions for future research.

6.1 Conclusions

The major achievement of this thesis is to show how news sentiment data can be used to improve the predictive ability of some existing time series' models that have useful roles in a variety of financial applications. The modified models offer a reasonable compromise between increased model complexity and prediction accuracy. Specific findings and contributions of different chapters are summarised below.

- In chapter 3, a large number of non-linear volatility models based on Gaussian processes combined with GARCH type models is constructed. Those different possible model structures are tested to determine whether the predictive ability of GARCH type models can be enhanced by using news sentiment scores. Particular emphasis has been given to the forecasting performance of the models and whether they can capture the characteristics of asset volatility and the influence of news on it. The empirical results of the suggested models showed that the prediction performances of the models vary across datasets and error measurement methods. We propose a novel model structure called *First* News Augmented GARCH (NA1-GARCH) model, which captures the stylistic features of the impact of news on stock volatility while preserving the familiar and parsimonious GARCH(1,1) form. The performance of NA1-GARCH model is compared with three other models, GARCH, EGARCH and TGARCH, each under two different distributional assumptions: Gaussian and student's t-distribution. We demonstrate through extensive numerical experiments, on 48 data sets drawn from two different stock markets, that the NA1-GARCH model outperforms the other three models in terms of out-of-sample volatility forecasting. The comparative analyses in evaluating the forecasting performance of the four models show that news sentiment has an impact on the daily volatility of stock market returns. In addition, the findings revealed that the same magnitude of positive (good news) and negative (bad news) shocks have different impact on the future volatility. The empirical results also showed that the student's t-distribution is more appropriate than the normal assumption, as it generates relatively more accurate forecasts. This analysis suggests that including the news impact term in the GARCH framework has improved the model prediction power for asymmetric type models. The empirical analysis later suggested that the NA1-GARCH model can be more useful than the other three models when estimating VaR and implement risk management strategies for FTSE100 and S&P500 stock index returns. Hence, the contribution of this research is to propose a model structure to incorporate a news sentiment term as an exogenous variable and to present extensive numerical evidence that including news sentiment significantly improves the predictive ability of GARCH type models.

- In chapter 4, the impact of high-frequency public news sentiment on the daily log returns volatility for 12 different stocks from two different stock markets (FTSE100 and EUROSTOXX50) was investigated. As in chapter 3, the quantified news sentiment and its impact on the movement of asset prices was considered as a second source of information, which is used together with the asset time series data to predict the volatility of asset price returns. A new volatility model was proposed, namely *Second* News Augmented GARCH (NA2-GARCH) model, which enables us to use the news sentiment score to improve the predictive ability of the GARCH model. The model structure of the NA2-GARCH model is chosen to be one with a direct multiplicative effect of news on the GARCH-predicted volatility. Numerical experiments indicated that an additive news impact model performs a lot worse than a multiplicative news impact model. This is another novel aspect of this work and a further contribution to knowledge. The computational results of the empirical investigation of more than 100 data sets drawn from the two sets of six assets are analysed. We then compared the performances of the GARCH, NA1-GARCH, NA2-GARCH and EGARCH models using the chosen performance measures. The results of the empirical experiments in this research clearly demonstrate that NA2-GARCH provides a superior prediction of volatility than the simple GARCH, NA1-GARCH and EGARCH models. The findings of this study also showed that positive news tends to reduce volatility whereas negative news tends to increase volatility. This analysis suggests that including the news impact term in the GARCH framework has improved the predictive ability of GARCH model. Another suggestion was that the use of an exponential decay function is good when the news flow is frequent, whereas the Hill decay function is good only when there are scheduled announcements, because the impact of the news on the market dies slowly as the impact of new information on a stock market depends on how unexpected the news is. The results vindicate the utility of our novel model structure combining the proxies for past news sentiments and the past asset price returns. NA2-GARCH is thus a computationally efficient means of exploiting the news sentiment score for better volatility prediction and it has a potential to be very useful in industrial practice. Therefore, the findings are crucial for all investors who are trading on the variance or volatility swaps, which can be used to speculate on future realized volatility, or to hedge the volatility exposure of other positions since the profit and loss from a variance swap depends directly on the difference between realized and implied volatility of a given underlying asset. In addition, NA2-GARCH model would be useful for investors who are focusing on risk-adjusted returns, especially those that utilize asset allocation and volatility targeting strategies. Furthermore, NA2-GARCH model can be used to estimate Value-at-Risk more accurately.

- In chapter 5, a group of one factor models for crude oil price modelling was proposed. The implementation of this group of models is based on the extension of an existing one factor with random mean model that was proposed by Islyaev (2014). This approach was extended by incorporating global macroeconomic news sentiment data. A method of incorporating macroeconomic news into a predictive model for forecasting prices of crude oil futures contracts was proposed. This study utilized the Kalman filtering framework for forecasting arbitrage-free (futures) prices, and assumed that the volatility of oil (futures) price is influenced by macroeconomic news. This study also used a one factor model with a constant risk premium, a random mean and a seasonality adjustment in terms of an additive sinusoid. The impact of quantified news sentiment on the price volatility is modelled through a parametrized, non-linear functional map. The news sentiment was used as an exogenous input which can change the volatility of the spot price. Specifically we used the global macroeconomic news sentiment applied to a broad dataset of the crude oil prices. We carried out empirical experiments for forecasting the futures contract prices of crude oil using the global macroeconomic news sentiment. The proposed model structure for incorporating macroeconomic news together with historical (market) data is novel and improves the accuracy of price prediction quite significantly. The first and the most important contribution of this study is that we demonstrate how to process macroeconomic news sentiment and how to incorporate it systematically into the commodity price model, in order to improve the forecasting of prices of crude oil futures. Our numerical experiments provide a strong evidence that macroeconomic news sentiment adds value to futures price forecasting. While most existing numerical studies only focus on short term futures contracts, our study includes the use of both short and long term futures contracts in model calibration as well as price forecasting. The results of this study has widespread uses for public and private sector entities which rely on crude oil, and also for economic policy makers who need to incorporate future oil price behaviour into their economic scenarios. The proposed methodology is clearly applicable for other commodities where futures prices are more liquid that the underlying commodity price, and where macroeconomic news data is available.

6.2 Future Research

The findings reported in this thesis can lead to several directions of research in the future, which can be of value to the financial modelling. Some of the possible extensions are listed below.

- This thesis focused on univariate models. More work can be done to generalise our findings in terms of multivariate predictive models.
- It is possible to use the news sentiment score as a switching signal in regime switching models, either in a filtering framework or in a GARCH type model. This could be related to using either filter with regime switching or GARCH models with switching. One can consider HMM-GARCH model and see if they can simplify significantly the 'hidden' part of HMM using News sentiment.
- One can also implement a particle filter to predict realized volatility from prices of derivative instruments which are non-linear in log spot price, and see if news sentiment would improve prediction of volatility.
- Investigating the impact of scheduled public company-specific announcements on asset price and volatility would require a somewhat different approach as

compared to aggregating the impact of unscheduled events as done in this thesis. Being able to use scheduled event sentiment in volatility forecasting will add further value to the toolkit being developed.

- It is also of interest to use news sentiment data as an additional source of information in spread modelling of commodity futures prices. The idea is to forecast the widening or shortening of spreads between two different futures of the same commodity, using the (linear) Kalman filter and news sentiment data.

Bibliography

Aiube, F. A. L., Baidya, T. K. N., and Tito, E. A. H. (2008). Analysis of commodity prices with the particle filter. *Energy Economics*, 30(2):597–605.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on Automatic Control*, 19(6):716–723.

Alberg, D., Shalit, H., and Yosef, R. (2008). Estimating stock market volatility using asymmetric GARCH models. *Applied Financial Economics*, 18(15):1201–1208.

Alquist, R. and Kilian, L. (2010). What do we learn from the price of crude oil futures? *Journal of Applied Econometrics*, 25(4):539–573.

Alquist, R., Kilian, L., and Vigfusson, R. J. (2013). Forecasting the price of oil. In *Handbook of economic forecasting*, volume 2, pages 427–507. Elsevier.

Andersen, T. G. (1996). Return volatility and trading volume: An information flow interpretation of stochastic volatility. *The Journal of Finance*, 51(1):169–204.

Andersen, T. G. and Bollerslev, T. (1997). Answering the critics: Yes, ARCH models do provide good volatility forecasts. Technical report, National Bureau of Economic Research.

Andersen, T. G. and Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International economic review*, 39(4):885–905.

Anderson, B. D. and Moore, J. B. (1979). Optimal filtering. *Englewood Cliffs*, 21:22–95.

Aragó, V. and Nieto, L. (2005). Heteroskedasticity in the returns of the main world stock exchange indices: volume versus GARCH effects. *Journal of International Financial Markets, Institutions and Money*, 15(3):271–284.

Arbex-Valle, C., Erlwein-Sayer, C., Kochendoerfer, A., Kuebler, B., Mitra, G., Nzouankeu-Nana, G.-A., Nouwt, B., and Stalknecht, B. (2013). News-enhanced market risk management. Available at SSRN: https://ssrn.com/abstract=2322668 or http://dx.doi.org/10.2139/ssrn.2322668.

Archanskaïa, E., Creel, J., and Hubert, P. (2012). The nature of oil shocks and the global economy. *Energy Policy*, 42:509–520.

Arshanapalli, B., d'Ouville, E., Fabozzi, F., and Switzer, L. (2006). Macroeconomic news effects on conditional volatilities in the bond and stock markets. *Applied Financial Economics*, 16(5):377–384.

Ashok, B., P. S. H. S. and Rahul, D. (2011). Impact of information arrival on volatility of intraday stock returns. *Working Paper*.

Avsar, S. G. and Goss, B. A. (2001). Forecast errors and efficiency in the US electricity futures market. *Australian Economic Papers*, 40(4):479–499.

Babbs, S. H. and Nowman, K. B. (1999). Kalman filtering of generalized Vasicek term structure models. *Journal of Financial and Quantitative Analysis*, 34(1):115–130.

Baillie, R. T. and Bollerslev, T. (2002). The message in daily exchange rates: a conditional-variance tale. *Journal of Business & Economic Statistics*, 20(1):60–68.

Baillie, R. T., Bollerslev, T., and Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 74(1):3–30.

Barndorff-Nielsen, O. E. and Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(2):253–280.

Baumeister, C. and Kilian, L. (2012). Real-time forecasts of the real price of oil. *Journal of Business & Economic Statistics*, 30(2):326–336.

Baumeister, C. and Peersman, G. (2013). The role of time-varying price elasticities in accounting for volatility changes in the crude oil market. *Journal of Applied Econometrics*, 28(7):1087–1109.

Black, F. (1976). Studies of stock price volatility changes. *In Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economics Section*, pages 177–181.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327.

Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *The review of economics and statistics*, 69(3):542–547.

Bollerslev, T. and Ghysels, E. (1996). Periodic autoregressive conditional heteroscedasticity. *Journal of Business & Economic Statistics*, 14(2):139–151.

Bozdogan, H. (2000). Akaike's information criterion and recent developments in information complexity. *Journal of mathematical psychology*, 44(1):62–91.

Brenner, R. J. and Kroner, K. F. (1995). Arbitrage, cointegration, and testing the unbiasedness hypothesis in financial markets. *Journal of Financial and Quantitative Analysis*, 30(1):23–42.

Burnham, K. and Anderson, D. (1998). *Model Selection and Inference: A Practical Information-theoretic Approach*. Intelligence, SS. of Lncs; 1501. Springer.

Burnham, K. P. and Anderson, D. R. (2003). *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Science & Business Media.

Campbell, S. D. (2005). A review of backtesting and backtesting procedures. Available at http://www.isihome.ir/freearticle/ISIHome.ir-24095.pdf.

Chen, X. and Ghysels, E. (2010). News -good or bad- and its impact on volatility predictions over multiple horizons. *The Review of Financial Studies*, 24(1):46–81.

Christoffersen, P. and Pelletier, D. (2004). Backtesting value-at-risk: A duration-based approach. *Journal of Financial Econometrics*, 2(1):84–108.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International economic review*, 39(4):841–862.

Cortazar, G., Milla, C., and Severino, F. (2008). A multicommodity model of futures prices: Using futures prices of one commodity to estimate the stochastic process of another. *Journal of Futures Markets*, 28(6):537–560.

Cortazar, G. and Schwartz, E. S. (2003). Implementing a stochastic model for oil futures prices. *Energy Economics*, 25(3):215–238.

Cousin, J.-G. and de Launois, T. (2006). New intensity and conditional volatility on the french stock market. *Finance*, 27(1):7–60.

Crouhy, M. and Rockinger, M. (1997). Volatility clustering, asymmetry and hysteresis in stock returns: International evidence. *Financial engineering and the Japanese markets*, 4(1):1–35.

Damodaran, A. (2007). *Strategic risk taking: a framework for risk management*. Pearson Prentice Hall.

Date, P. and Ponomareva, K. (2011). Linear and non-linear filtering in mathematical finance: a review. *IMA Journal of Management Mathematics*, 22(3):195–211.

Date, P. and Wang, C. (2009). Linear gaussian affine term structure models with unobservable factors: Calibration and yield forecasting. *European Journal of Operational Research*, 195(1):156–166.

Dowd, K. (1998). *Beyond value at risk: the new science of risk management*. John Wiley & Sons.

Dowd, K. (2006). Retrospective assessment of value at risk. In *Risk Management*, pages 183–202. Elsevier.

Duffie, D. (2010). *Dynamic asset pricing theory*. Princeton University Press.

Elder, J., Miao, H., and Ramchander, S. (2012). Impact of macroeconomic news on metal futures. *Journal of Banking and Finance*, 36(1):51 – 65.

Emery, G. W. and Liu, Q. W. (2002). An analysis of the relationship between electricity and natural-gas futures prices. *Journal of Futures Markets*, 22(2):95–122.

Engel, R. (1990). Discussion: stock market volatility and the crash. *Review of Financial Studies*, 3:103–106.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, 50(4):987–1007.

Engle, R. F. and Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *The journal of finance*, 48(5):1749–1778.

Engle, R. F., Patton, A. J., et al. (2001). What good is a volatility model. *Quantitative finance*, 1(2):237–245.

Erlwein-Sayer, C. (2017). Forecasting sovereign bond spreads with macroeconomic news sentiment. *Working paper*. Available at: http://www.ps-quant.com/SR/wp-content/uploads/2018/01/WhitePaper_ForecastingSovereignBondSpreads.pdf.

Feng, Y., Beran, J., and Yu, K. (2007). Modelling financial time series with SEMIFAR-GARCH model. *IMA Journal of Management Mathematics*, 18(4):395–412.

Finger, C. (2006). How historical simulation made me lazy. *RiskMetrics Group Research Monthly*.

Franses, P. H. and Van Dijk, D. (1996). Forecasting stock market volatility using (non-linear) GARCH models. *Journal of Forecasting*, 15(3):229–235.

French, K. R., Schwert, G. W., and Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of financial Economics*, 19(1):3–29.

Fuller, W. (1976). *Introduction to statistical time series*. New York: John Wiley & Sons.

Galati, G. and Ho, C. (2003). Macroeconomic news and the Euro/Dollar exchange rate. *Economic Notes*, 32(3):371–398.

Girma, P. B. and Paulson, A. S. (1999). Risk arbitrage opportunities in petroleum futures spreads. *Journal of Futures Markets*, 19(8):931–955.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, 48(5):1779–1801.

Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, 42(1):27–62.

Hafez, P. (2009). Impact of news sentiment on abnormal stock returns. *White paper, RavenPack*.

Hafez, P. (2013). Trading relative value based on news indicators. *White paper, RavenPack*.

Hall, A. R. (2005). *Generalized method of moments*. Oxford University Press.

Hansen, P. R. and Lunde, A. (2005). A forecast comparison of volatility models: does anything beat a GARCH (1, 1)? *Journal of applied econometrics*, 20(7):873–889.

Hess, D., Huang, H., and Niessen, A. (2008). How do commodity futures respond to macroeconomic news? *Financial Markets and Portfolio Management*, 22(2):127–146.

Hien, M. T. T. and Thanh, T. (2008). Modelling and forecasting volatility by GARCH-type models: the case of vietnam stock exchange. *A dissertation presented in part consideration for the degree of MA. Finance and Investment*, pages 1–97.

Hill, A. V. (1910). The possible effects of the aggregation of the molecules of hemoglobin on its dissociation curves. *Journal of Physiology*, 40:iv–vii.

Ho, K.-Y., Shi, Y., and Zhang, Z. (2013). How does news sentiment impact asset volatility? evidence from long memory and regime-switching approaches. *The North American Journal of Economics and Finance*, 26:436–456.

Hsieh, D. A. (1988). The statistical properties of daily foreign exchange rates: 1974–1983. *Journal of international economics*, 24(1-2):129–145.

Hsieh, D. A. (1989). Modeling heteroscedasticity in daily foreign-exchange rates. *Journal of Business & Economic Statistics*, 7(3):307–317.

Hsu, T.-K., Tsai, C.-C., and Cheng, K.-L. (2016). Forecast of 2013–2025 crude oil prices: Quadratic sine-curve trend model application. *Energy Sources, Part B: Economics, Planning, and Policy*, 11(3):205–211.

Hyndman, C. B. (2007). Gaussian factor models-futures and forward prices. *IMA Journal of Management Mathematics*, 18(4):353–369.

Islyaev, S. (2014). *Stochastic models with random parameters for financial markets*. PhD thesis, Brunel University London.

Jorion, P. (1988). On jump processes in the foreign exchange and stock markets. *The Review of Financial Studies*, 1(4):427–445.

Jorion, P. (1997). *Value at risk: the new benchmark for controlling market risk*. Irwin Professional Pub.

Jorion, P. (2001). Value at risk: the new benchmark for managing financial risk. *NY: McGraw-Hill Professional*.

Kalev, P. S., Liu, W.-M., Pham, P. K., and Jarnecic, E. (2004). Public information arrival and volatility of intraday stock returns. *Journal of Banking & Finance*, 28(6):1441–1467.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45.

Kilian, L. (2009). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review*, 99(3):1053–69.

Kilian, L. and Murphy, D. P. (2012). Why agnostic sign restrictions are not enough: understanding the dynamics of oil market VaR models. *Journal of the European Economic Association*, 10(5):1166–1188.

Kilian, L. and Murphy, D. P. (2014). The role of inventories and speculative trading in the global market for crude oil. *Journal of Applied Econometrics*, 29(3):454–478.

Kim, S.-J., McKenzie, M. D., and Faff, R. W. (2004). Macroeconomic news announcements and the role of expectations: evidence for US bond, stock and foreign exchange markets. *Journal of Multinational Financial Management*, 14(3):217 – 232.

Kritzman, M. and Rich, D. (2002). The mismeasurement of risk. *Financial Analysts Journal*, 58(3):91–99.

Kumar, M. S. (1992). The forecasting accuracy of crude oil futures prices. *IMF Economic Review*, 39(2):432–461.

Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *The journal of Derivatives*, 3(2):73–84.

Lautier, D. and Galli, A. (2004). Simple and extended Kalman filters: an application to term structures of commodity prices. *Applied Financial Economics*, 14(13):963–973.

Lee, G. and Engle, R. (1993). A permanent and transitory component model of stock return volatility. Available at SSRN: https://ssrn.com/abstract=5848.

Li, L. and Engle, R. F. (1998). Macroeconomic announcements and volatility of treasury futures. *Department of Economics, UCSD*.

Linsmeier, T. J., Pearson, N. D., et al. (1996). Risk measurement: An introduction to value at risk. Available at: https://www.casact.org/education/specsem/99frmgt/pearson2.pdf.

Liu, H.-C., Lee, Y.-H., and Lee, M.-C. (2009). Forecasting china stock markets volatility via GARCH models under skewed-GED distribution. *Journal of money, Investment and Banking*, 7(1):5–15.

Longin, F. M. (2001). Beyond the VaR. *The Journal of Derivatives*, 8(4):36–48.

Lütkebohmert, E. (2008). *Concentration risk in credit portfolios*. Springer Science & Business Media.

Maheu, J. M. and McCurdy, T. H. (2004). News arrival, jump dynamics, and volatility components for individual stock returns. *The Journal of Finance*, 59(2):755–793.

Manoliu, M. and Tompaidis, S. (2002). Energy futures prices: term structure models with Kalman filter estimation. *Applied Mathematical Finance*, 9(1):21–43.

McCallum, A. H. and Wu, T. (2005). Do oil futures prices help predict future oil prices? *Economic Letter*, 38. Available at SSRN: https://ssrn.com/abstract=1967966 or http://dx.doi.org/10.2139/ssrn.1967966.

McCurdy, T. H. and Morgan, I. G. (1987). Tests of the martingale hypothesis for foreign currency futures with time-varying volatility. *International Journal of Forecasting*, 3(1):131–148.

Mirantes, A. G., PoblaciÃşn, J., and Serna, G. (2012). The stochastic seasonal behaviour of natural gas prices. *European Financial Management*, 18(3):410–443.

Mitchell, M. L. and Mulherin, J. H. (1994). The impact of public information on the stock market. *The Journal of Finance*, 49(3):923–950.

Mitra, G. and Mitra, L. (2011). *The handbook of news analytics in finance*, volume 596. John Wiley & Sons.

Mitra, L. R., Mitra, G., and di Bartolomeo, D. (2008). Equity portfolio risk (volatility) estimation using market information and sentiment. *Quantitative Finance*, 9(8):887–895.

Monoyios, M. (2007). Optimal hedging and parameter uncertainty. *IMA Journal of Management Mathematics*, 18(4):331–351.

Moon, K. (1997). A understanding of vector autoregressive model. *J. Korean Off. Stat*, 2:23–56.

Moosa, I. A. and Al-Loughani, N. E. (1994). Unbiasedness and time varying risk premia in the crude oil futures market. *Energy economics*, 16(2):99–105.

Morana, C. (2013). The oil price-macroeconomy relationship since the mid-1980s: A global perspective. *The Energy Journal*, 34(3):153–189.

Moshiri, S. and Foroutan, F. (2006). Forecasting nonlinear crude oil futures prices. *The Energy Journal*, 27(4):81–95.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 59(2):347–370.

Nikkinen, J., Omran, M., and ahlstr P. (2006). Global stock market reactions to scheduled U.S. macroeconomic news announcements. *Global Finance Journal*, 17(1):92 – 104.

Øksendal, B. K. (2003). *Stochastic Differential Equations: An Introduction with Applications,*. Springer, Heidelberg.

Pagano, P. and Pisani, M. (2009). Risk-adjusted forecasts of oil prices. *The BE Journal of Macroeconomics*, 9(1).

RavenPack (2014). Ravenpack news analytics, version 4.0: User guide and service overview.

Reeve, T. A. and Vigfusson, R. (2011). Evaluating the forecasting performance of commodity futures prices. 1025. Available at SSRN: https://ssrn.com/abstract=1912969 or http://dx.doi.org/10.2139/ssrn.1912969.

Reuters, T. (2010). Thomson reuters news analytics. version 2.0.2.

Riordan, R., Storkenmaier, A., Wagener, M., and Zhang, S. S. (2013). Public information arrival: Price discovery and liquidity in electronic limit order markets. *Journal of Banking & Finance*, 37(4):1148–1159.

Schwartz, E. S. (1997). The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of Finance*, 52(3):923–973.

Schwarz, G. et al. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464.

Sharma, J. L., Mougoue, M., and Kamath, R. (1996). Heteroscedasticity in stock market indicator return data: volume versus GARCH effects. *Applied Financial Economics*, 6(4):337–342.

Sidorov, S., Date, P., Balash, V., et al. (2013). Using news analytics data in GARCH models. *Applied Econometrics*, 29(1):82–96.

Sidorov, S. P., Date, P., and Balash, V. (2014). GARCH type volatility models augmented with news intensity data. In *Chaos, Complexity and Leadership 2012*, pages 199–207. Springer.

Silvapulle, P. and Moosa, I. A. (1999). The relationship between spot and futures prices: evidence from the crude oil market. *Journal of Futures Markets*, 19(2):175–193.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: Journal of the Econometric Society*, 48(1):1–48.

Song, J.-Y. (2010). Asymmetric impact of news on stock return volatility. In *Banking And Capital Markets: New International Perspectives*, pages 373–409. World Scientific.

Supervision, B. (2011). Basel committee on banking supervision. Available at https://www.bis.org/publ/bcbs213.pdf, December 2011.

Taylor, S. J. (1986). *Modelling financial time series*. Chichester.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3):1139–1168.

Tetlock, P. C. (2010). Does public financial news resolve asymmetric information? *The Review of Financial Studies*, 23(9):3520–3557.

Timilsina, G. R. (2015). Oil prices and the global economy: A general equilibrium analysis. *Energy Economics*, 49:669–675.

Tsai, K.-T. (2004). Risk management via value at risk. *ICSA Bulletin*, pages 14–23.

Tsay, R. (2008). Out-of-sample forecasts. *Lecture Notes, The University of Chicago Booth School of Business.* http://faculty.chicagobooth.edu/ruey.tsay/teaching/bs41202/sp2008/lec4a-08.pdf. Accessed January 2017.

Tsay, R. S. (2005). *Analysis of financial time series*, volume 543. John Wiley & Sons.

Weiss, N. A. and Hassett, M. J. (1999). *Introductory statistics*. Addison-Wesley Boston.

Xiao, J., Brooks, R. D., and Wong, W.-K. (2009). GARCH and volume effects in the australian stock markets. *Annals of Financial Economics*, 5(01):0950005.

Yu, X. (2014). *Analysis of news sentiment and its applications to finance*. PhD thesis, Brunel University London, London - UK.

Yu, X., Mitra, G., Arbex-Valle, C., and Sayer, T. (2015). An impact measure for news: Its use in daily trading strategies. Available at SSRN: https://ssrn.com/abstract=2702032 or http://dx.doi.org/10.2139/ssrn.2702032.

Zakoian, J.-M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and control*, 18(5):931–955.

Zellner, A. (1976). Bayesian and non-Bayesian analysis of the regression model with multivariate Student-t error terms. *Journal of the American Statistical Association*, 71(354):400–405.

Appendix A

Chapter 3

A.1 Tables of parameter estimates

Datasets	GARCH			NA1-GARCH						TGARCH				EGARCH			
	ω	α	β	ω	α	β	κ	γ	λ	ω	α^+	α^-	β	ω	α	β	γ
FTSE1	1.99E-06	9.45E-02	8.60E-01	1.65E-06	1.12E-01	8.88E-01	5.16E-02	5.66E-02	1.02E+00	1.01E-13	6.08E-02	2.95E-01	8.85E-01	4.84E-06	1.52E-01	9.99E-01	-8.45E-02
FTSE2	2.53E-06	9.25E-02	8.55E-01	1.92E-06	1.11E-01	8.46E-01	6.53E-02	7.86E-02	9.80E-01	5.52E-12	2.35E-08	2.30E-02	9.96E-01	1.24E-06	1.53E-01	9.99E-01	-9.23E-02
FTSE3	3.26E-06	9.74E-02	8.39E-01	1.97E-06	1.17E-01	8.34E-01	4.90E-02	1.05E-01	9.71E-01	4.65E-05	5.13E-08	1.18E-01	9.53E-01	-1.20E-05	1.42E-01	9.99E-01	-9.97E-02
FTSE4	3.88E-06	1.54E-01	8.02E-01	2.30E-06	9.01E-02	3.97E-01	3.18E-02	2.84E-02	5.07E-01	1.58E-04	2.82E-07	1.54E-01	9.24E-01	9.16E-07	1.77E-01	9.99E-01	-1.05E-01
FTSE5	3.04E-06	1.54E-01	8.30E-01	1.61E-06	1.06E-01	5.65E-01	8.20E-02	1.25E-01	6.63E-01	3.42E-04	2.54E-06	1.87E-01	8.94E-01	1.67E-06	1.65E-01	9.99E-01	-1.24E-01
FTSE6	2.40E-06	1.29E-01	8.62E-01	1.74E-06	1.07E-01	6.65E-01	7.10E-02	1.36E-01	7.64E-01	4.43E-16	7.84E-03	3.17E-04	1.01E+00	-2.13E-05	1.11E-01	9.99E-01	-9.01E-02
FTSE7	4.34E-06	1.31E-01	8.69E-01	6.86E-07	1.09E-01	8.91E-01	1.13E-01	1.64E-01	9.67E-01	6.53E-05	8.82E-08	1.28E-01	9.46E-01	-1.20E-05	1.37E-01	1.00E+00	-1.10E-01
FTSE8	3.60E-06	1.09E-01	8.91E-01	3.06E-06	6.55E-02	5.16E-01	3.89E-02	5.56E-06	5.75E-01	1.00E-05	2.52E-10	1.17E-01	9.53E-01	8.64E-06	9.77E-02	1.00E+00	-1.16E-01
FTSE9	6.72E-06	9.79E-02	8.84E-01	1.25E-06	9.91E-02	9.01E-01	1.47E-02	1.60E-01	9.76E-01	2.55E-09	4.84E-12	7.19E-03	1.00E+00	4.50E-07	1.35E-01	1.00E+00	-9.33E-02
FTSE10	3.50E-06	9.22E-02	8.97E-01	1.21E-06	1.08E-01	8.59E-01	9.16E-03	1.52E-01	9.43E-01	3.44E-04	2.27E-04	1.33E-01	9.27E-01	1.15E-06	1.37E-01	9.99E-01	-8.22E-02
FTSE11	2.75E-06	1.17E-01	8.81E-01	7.34E-07	1.34E-01	8.66E-01	1.84E-02	1.64E-01	9.66E-01	2.91E-06	1.72E-01	6.87E-01	7.61E-01	8.17E-07	1.94E-01	9.99E-01	-1.15E-01
FTSE12	3.69E-06	5.71E-02	9.18E-01	1.44E-06	7.71E-02	7.85E-01	8.87E-02	1.39E-01	8.49E-01	1.44E-14	3.37E-04	1.13E-02	9.97E-01	-8.41E-03	1.81E-01	9.98E-01	-1.09E-01
FTSE13	3.25E-06	6.67E-02	9.11E-01	2.68E-06	4.02E-02	4.57E-01	4.55E-02	3.76E-02	5.07E-01	2.89E-06	3.33E-02	1.36E-01	9.39E-01	-4.85E-07	3.99E-01	9.97E-01	-7.99E-02
FTSE14	5.89E-06	7.37E-02	8.72E-01	2.27E-06	4.94E-02	5.72E-01	7.34E-02	1.77E-01	6.22E-01	5.28E-04	3.53E-11	1.87E-01	8.80E-01	-3.02E-07	1.47E-01	9.99E-01	-1.14E-01
FTSE15	1.39E-05	9.91E-02	7.70E-01	3.26E-06	3.01E-02	2.31E-01	7.05E-02	1.69E-01	2.73E-01	1.96E-14	1.72E-11	2.62E-02	1.00E+00	-9.35E-04	1.10E-01	9.99E-01	-1.07E-01
FTSE16	4.99E-06	1.16E-01	8.50E-01	2.24E-06	6.85E-02	5.31E-01	4.19E-02	1.84E-01	5.90E-01	2.55E-04	2.08E-03	1.73E-01	9.10E-01	6.63E-07	1.61E-01	9.99E-01	-1.26E-01
FTSE17	3.85E-06	7.64E-02	8.90E-01	1.54E-06	6.00E-02	7.31E-01	9.19E-02	2.16E-01	7.77E-01	9.45E-13	2.14E-09	1.33E-02	9.97E-01	2.62E-06	1.17E-01	1.00E+00	-1.33E-01
FTSE18	3.68E-06	8.32E-02	8.86E-01	1.34E-06	7.63E-02	8.33E-01	3.65E-02	1.52E-01	8.94E-01	9.80E-05	3.07E-11	1.31E-01	9.40E-01	1.52E-06	5.70E-02	1.00E+00	-1.54E-01
FTSE19	2.26E-06	7.53E-02	9.04E-01	2.32E-06	5.51E-02	5.43E-01	3.26E-04	7.41E-02	6.01E-01	4.16E-16	6.95E-04	3.46E-02	9.92E-01	-2.72E-06	-5.79E-02	1.00E+00	-1.20E-01
FTSE20	2.53E-06	8.31E-02	8.93E-01	2.02E-06	8.78E-02	5.51E-01	3.09E-02	2.26E-01	6.04E-01	4.16E-05	5.86E-12	5.34E-02	9.75E-01	-9.06E-07	1.79E-01	9.99E-01	-7.39E-02
FTSE21	2.80E-06	4.11E-02	9.18E-01	1.09E-06	5.75E-02	9.20E-01	1.28E-01	1.92E-01	9.62E-01	1.12E-05	2.56E-10	1.61E-02	9.92E-01	2.64E-07	8.75E-02	9.99E-01	-3.48E-02
FTSE22	1.19E-06	3.18E-02	9.48E-01	2.74E-06	1.04E-01	6.55E-01	3.67E-02	1.09E-06	7.77E-01	1.17E-05	2.24E-02	4.83E-02	9.72E-01	-3.83E-05	7.81E-02	1.00E+00	-2.73E-02
FTSE23	5.83E-06	7.95E-02	8.10E-01	1.83E-06	1.15E-01	8.75E-01	1.33E-01	6.57E-02	9.92E-01	5.61E-04	3.99E-12	1.34E-01	8.73E-01	2.82E-06	7.78E-02	1.00E+00	-3.01E-02
FTSE24	3.17E-06	7.50E-02	8.62E-01	1.57E-06	8.13E-02	7.21E-01	1.17E-01	1.33E-01	8.07E-01	5.27E-16	3.31E-01	8.82E-01	6.95E-01	-1.12E-04	1.09E-01	9.99E-01	-4.62E-02

Table A.1: Parameter estimates of GARCH(1,1), NA1-GARCH(1,1), TGARCH(1,1) and EGARCH(1,1) models for FTSE100 datasets (Normal distribution)

Datasets	GARCH			NA1-GARCH						TGARCH				EGARCH			
	ω	α	β	ω	α	β	κ	γ	λ	ω	α^+	α^-	β	ω	α	β	γ
SP1	2.01E-06	4.72E-02	9.03E-01	2.53E-06	1.21E-01	8.39E-01	1.99E-01	1.36E-07	9.77E-01	5.13E-13	1.54E-01	1.96E-01	8.75E-01	-2.78E-05	-4.42E-02	1.00E+00	-6.70E-02
SP2	2.36E-06	3.10E-02	9.12E-01	1.91E-06	8.02E-02	9.14E-01	9.03E-02	1.48E-01	1.01E+00	1.02E-08	5.55E+00	7.72E-01	3.37E-01	5.03E-06	-2.01E-02	1.00E+00	-5.88E-02
SP3	2.47E-06	3.80E-02	9.10E-01	1.39E-06	5.36E-02	8.02E-01	1.89E-01	2.68E-01	8.54E-01	5.04E-13	1.66E-02	2.60E-06	9.99E-01	-6.54E-07	7.35E-02	9.99E-01	-7.06E-02
SP4	1.12E-06	5.29E-02	9.35E-01	1.35E-06	6.91E-02	7.82E-01	1.42E-01	1.59E-01	8.45E-01	2.36E-05	1.04E-03	6.76E-02	9.75E-01	5.98E-06	-6.24E-02	1.00E+00	-1.38E-01
SP5	1.08E-06	5.50E-02	9.37E-01	2.72E-06	8.22E-02	8.17E-01	3.79E-02	2.99E-02	9.20E-01	3.95E-13	9.24E-10	2.06E-01	9.40E-01	1.62E-06	7.29E-02	9.99E-01	-5.50E-02
SP6	3.63E-07	8.50E-03	9.91E-01	1.85E-06	4.91E-02	5.27E-01	9.76E-02	6.82E-02	5.78E-01	2.73E-17	2.35E-02	4.09E-04	9.95E-01	-1.07E-06	6.38E-02	9.99E-01	-2.14E-02
SP7	3.97E-06	9.41E-02	9.00E-01	8.93E-07	8.73E-02	9.11E-01	2.84E-01	1.14E-01	9.50E-01	2.70E-14	1.57E-02	3.64E-04	1.00E+00	6.01E-07	1.62E-01	1.00E+00	-1.13E-01
SP8	2.55E-06	8.91E-02	9.11E-01	1.36E-06	6.08E-02	5.27E-01	3.29E-01	8.44E-02	5.31E-01	3.44E-16	3.68E-10	1.63E-02	1.00E+00	-2.58E-06	1.34E-01	1.00E+00	-1.23E-01
SP9	4.34E-06	9.16E-02	8.98E-01	9.09E-07	1.09E-01	8.91E-01	2.26E-01	1.35E-01	9.59E-01	3.12E-04	2.82E-08	1.50E-01	9.23E-01	-3.85E-06	1.50E-01	1.00E+00	-1.09E-01
SP10	3.06E-06	9.91E-02	8.94E-01	8.33E-07	8.69E-02	9.13E-01	3.11E-01	2.03E-03	9.26E-01	3.64E-16	5.71E-07	3.20E-02	9.90E-01	-9.34E-06	1.62E-01	1.00E+00	-1.19E-01
SP11	2.36E-06	1.18E-01	8.82E-01	2.50E-06	9.11E-02	6.73E-01	1.16E-01	3.46E-02	7.47E-01	2.84E-14	6.95E-02	1.45E-01	9.21E-01	-1.39E-06	2.31E-01	9.99E-01	-8.88E-02
SP12	2.82E-06	6.99E-02	9.12E-01	9.97E-07	9.22E-02	7.01E-01	1.36E-01	1.50E-01	7.73E-01	4.41E-14	7.72E-10	1.75E-02	9.93E-01	4.15E-07	1.97E-01	9.99E-01	-7.28E-02
SP13	1.98E-06	7.37E-02	9.11E-01	2.86E-06	7.22E-02	5.53E-01	3.35E-02	9.94E-03	6.29E-01	2.10E-17	9.89E-11	8.49E-03	1.00E+00	2.07E-06	1.50E-01	9.99E-01	-1.06E-01
SP14	2.89E-06	8.25E-02	8.91E-01	2.85E-06	6.34E-02	5.67E-01	3.92E-03	2.98E-03	6.52E-01	6.22E-18	2.66E-02	3.59E-04	9.91E-01	-5.86E-06	6.30E-01	9.91E-01	1.08E-01
SP15	3.85E-06	1.32E-01	8.49E-01	1.72E-06	9.16E-02	7.14E-01	7.95E-02	9.76E-02	7.97E-01	4.46E-04	4.44E-06	1.57E-01	9.02E-01	-3.76E-06	2.22E-01	9.98E-01	-8.17E-02
SP16	3.42E-06	1.28E-01	8.53E-01	8.82E-07	4.68E-02	4.49E-01	1.34E-01	1.69E-01	4.80E-01	3.47E-04	5.33E-02	2.17E-01	8.75E-01	3.96E-08	2.42E-01	9.98E-01	-7.62E-02
SP17	2.85E-06	9.94E-02	8.78E-01	1.52E-06	7.20E-02	7.30E-01	1.23E-01	1.44E-01	7.95E-01	1.58E-12	2.29E-06	6.37E-02	9.78E-01	2.23E-06	-1.01E-01	1.01E+00	2.14E-01
SP18	3.88E-06	1.22E-01	8.46E-01	1.58E-06	8.98E-02	6.74E-01	1.13E-01	1.33E-01	7.54E-01	1.16E-15	4.75E-02	1.63E-08	9.86E-01	-2.67E-07	1.87E-01	9.98E-01	-1.04E-01
SP19	4.28E-06	1.50E-01	8.19E-01	1.60E-06	5.77E-02	4.52E-01	8.63E-02	9.20E-02	5.09E-01	2.55E-03	1.49E-01	5.43E-01	5.46E-01	-3.95E-05	2.05E-01	9.98E-01	-1.45E-01
SP20	4.05E-06	1.47E-01	8.23E-01	1.16E-06	8.43E-02	7.04E-01	2.22E-01	2.32E-02	7.12E-01	6.74E-04	2.59E-03	3.54E-01	8.16E-01	1.00E-06	9.00E-02	9.00E-01	5.90E-01
SP21	4.36E-06	5.01E-02	8.77E-01	1.44E-06	6.77E-02	7.19E-01	1.63E-01	7.54E-02	7.68E-01	6.02E-06	1.28E-04	7.26E-08	1.00E+00	1.18E-06	1.46E-01	9.99E-01	-1.27E-01
SP22	8.96E-06	9.38E-02	7.57E-01	3.12E-06	3.37E-02	5.78E-01	2.41E-06	2.41E-03	6.64E-01	1.91E-07	1.68E+00	2.75E-01	6.21E-01	-1.06E-06	8.49E-02	9.99E-01	-1.12E-01
SP23	1.15E-05	1.35E-01	6.43E-01	1.86E-06	9.92E-02	7.92E-01	6.55E-02	8.57E-02	9.10E-01	7.88E-08	1.13E-02	1.99E-03	9.99E-01	1.53E-06	-5.33E-02	1.00E+00	-8.88E-02
SP24	9.12E-06	1.96E-01	6.24E-01	2.83E-06	9.35E-02	3.34E-01	9.01E-09	1.12E-01	4.59E-01	8.70E-04	4.68E-09	2.83E-01	7.78E-01	-2.38E-05	2.11E-01	9.98E-01	-1.75E-01

Table A.2: Parameter estimates of GARCH(1,1), TGARCH(1,1) and N-GARCH(1,1) models for SP500 datasets (Normal distribution)

Datasets	GARCH				NA1-GARCH							TGARCH			EGARCH						
	ω	α	β	*df*	ω	α	β	κ	γ	λ	*df*	ω	α^+	α^-	β	*df*	ω	α	β	γ	*df*
FTSE1	2.01E-06	9.42E-02	8.60E-01	198	5.97E-14	9.43E-02	8.89E-01	7.32E-02	1.92E-01	9.57E-01	15	4.59E-07	7.17E-10	3.74E-02	9.90E-01	6	-3.97E-06	1.49E-01	9.99E-01	-8.57E-02	343
FTSE2	1.98E-06	8.60E-02	8.73E-01	198	6.32E-07	1.28E-01	8.72E-01	9.74E-02	1.23E-01	9.79E-01	16	2.27E-07	4.52E-05	4.30E-02	9.86E-01	10	-1.56E-05	1.60E-01	9.99E-01	-9.58E-02	343
FTSE3	3.26E-06	9.84E-02	8.38E-01	198	2.04E-06	1.02E-01	7.83E-01	9.30E-02	5.02E-04	8.99E-01	95	8.80E-16	3.26E-02	6.21E-12	9.89E-01	8	8.25E-07	1.43E-01	9.99E-01	-9.96E-02	86
FTSE4	3.86E-06	1.54E-01	8.02E-01	198	2.19E-06	1.14E-01	6.78E-01	4.87E-02	1.36E-01	8.02E-01	98	1.94E-09	2.54E-02	1.14E-11	9.93E-01	6	-1.59E-06	1.71E-01	9.99E-01	-1.10E-01	343
FTSE5	3.04E-06	1.53E-01	8.30E-01	198	1.56E-06	1.19E-01	8.81E-01	1.21E-01	4.76E-02	9.84E-01	71	7.03E-07	3.79E-02	3.36E-04	9.89E-01	6	-2.05E-06	1.72E-01	9.99E-01	-1.22E-01	343
FTSE6	2.41E-06	1.27E-01	8.62E-01	198	1.81E-06	9.69E-02	9.02E-01	1.02E-01	8.89E-02	9.91E-01	79	4.80E-06	1.77E-01	2.23E-01	8.63E-01	11	1.31E-06	1.47E-01	9.99E-01	2.08E-02	115
FTSE7	1.28E-06	9.96E-02	9.00E-01	198	5.56E-07	1.29E-01	7.85E-01	6.81E-02	2.15E-01	8.80E-01	78	6.72E-08	2.85E-02	8.68E-03	9.94E-01	9	1.69E-06	1.45E-01	1.00E+00	-1.13E-01	343
FTSE8	8.18E-06	1.20E-01	8.64E-01	198	4.30E-07	1.20E-01	8.55E-01	8.66E-02	2.71E-01	9.36E-01	90	3.05E-05	1.55E-02	1.84E-04	9.96E-01	9	2.52E-06	9.02E-02	1.00E+00	-1.08E-01	129
FTSE9	4.08E-06	1.08E-01	8.90E-01	198	3.44E-06	1.28E-01	8.72E-01	1.63E-02	1.68E-02	9.95E-01	94	1.92E-06	9.19E-05	2.23E-02	9.92E-01	6	3.85E-07	1.38E-01	9.99E-01	-9.14E-02	343
FTSE10	3.08E-06	9.44E-02	8.98E-01	198	2.75E-10	1.28E-01	7.36E-01	1.68E-01	4.45E-02	8.03E-01	56	2.21E-07	1.48E-02	1.13E-01	9.48E-01	64	9.84E-07	1.48E-01	9.99E-01	-8.50E-02	343
FTSE11	1.60E-06	1.07E-01	8.93E-01	169	1.56E-06	1.07E-01	8.04E-01	1.55E-01	1.42E-01	8.81E-01	57	7.89E-13	2.34E-02	7.50E-03	9.88E-01	5	1.58E-06	1.92E-01	9.99E-01	-1.13E-01	343
FTSE12	3.53E-06	5.42E-02	9.21E-01	198	6.79E-08	1.45E-01	8.55E-01	1.65E-01	7.86E-02	9.51E-01	49	2.82E-05	1.77E-01	4.96E-01	7.89E-01	33	-1.26E-05	1.59E-01	1.00E+00	-9.47E-02	342
FTSE13	3.21E-06	6.54E-02	9.12E-01	198	2.85E-06	4.62E-02	7.48E-01	6.71E-02	5.17E-02	8.02E-01	76	8.11E-06	4.24E-04	4.40E-02	9.84E-01	7	-6.84E-06	1.35E-01	9.99E-01	-9.45E-02	343
FTSE14	2.07E-06	6.21E-02	9.19E-01	198	3.20E-06	3.51E-02	5.60E-01	2.76E-02	1.31E-01	6.09E-01	57	4.32E-05	1.26E-01	3.44E-01	8.46E-01	195	-1.89E-06	1.58E-01	9.99E-01	-1.18E-01	343
FTSE15	1.39E-05	9.76E-02	7.71E-01	198	2.71E-06	3.96E-02	5.10E-01	7.58E-02	2.27E-01	5.50E-01	52	1.66E-04	1.58E-01	5.54E-01	7.66E-01	16	-2.23E-06	4.24E-01	9.95E-01	-2.25E-01	343
FTSE16	4.96E-06	1.15E-01	8.51E-01	198	2.28E-06	1.02E-01	7.20E-01	7.93E-02	3.24E-03	8.12E-01	66	4.65E-09	2.03E-02	3.06E-02	9.83E-01	7	-5.49E-06	1.81E-01	9.99E-01	-1.27E-01	320
FTSE17	3.83E-06	7.65E-02	8.90E-01	198	1.21E-06	6.11E-02	8.28E-01	7.32E-02	1.93E-01	8.75E-01	148	1.87E-05	9.61E-02	2.70E-01	8.70E-01	72	-4.35E-06	1.14E-01	1.00E+00	-1.33E-01	343
FTSE18	3.68E-06	8.30E-02	8.86E-01	198	2.03E-06	6.20E-02	7.97E-01	8.97E-02	9.15E-02	8.59E-01	87	3.63E-09	9.98E-03	1.73E-02	9.92E-01	5	-2.48E-06	5.23E-02	1.00E+00	-1.48E-01	343
FTSE19	2.22E-06	7.57E-02	9.04E-01	198	1.88E-06	7.51E-02	7.74E-01	1.12E-01	9.58E-02	8.44E-01	97	8.08E-07	1.60E-06	3.04E-02	9.88E-01	6	1.66E-06	-5.84E-02	1.00E+00	-1.51E-01	143
FTSE20	2.45E-06	8.28E-02	8.94E-01	198	1.62E-06	8.92E-02	9.07E-01	1.18E-01	9.63E-02	9.90E-01	71	2.17E-06	1.04E-04	2.76E-02	9.89E-01	6	-6.47E-06	1.51E-01	9.99E-01	-6.91E-02	343
FTSE21	2.74E-06	4.08E-02	9.19E-01	195	1.49E-06	1.08E-01	8.92E-01	1.34E-01	1.72E-01	9.86E-01	87	3.55E-06	1.60E-01	5.78E-02	9.23E-01	198	1.49E-05	8.35E-02	1.00E+00	-3.40E-02	189
FTSE22	1.48E-06	3.81E-02	9.38E-01	198	2.36E-06	5.34E-02	7.17E-01	5.40E-02	6.49E-02	7.99E-01	125	2.26E-05	1.96E-02	3.01E-01	8.98E-01	12	2.20E-06	1.84E-01	9.98E-01	-5.93E-02	343
FTSE23	2.72E-06	5.07E-02	8.95E-01	198	2.72E-06	7.19E-02	5.48E-01	9.93E-05	1.07E-01	6.41E-01	58	8.01E-11	1.95E-02	4.72E-03	9.92E-01	6	-5.96E-07	8.49E-02	9.99E-01	-3.49E-02	343
FTSE24	3.15E-06	7.52E-02	8.62E-01	198	1.46E-06	1.54E-01	8.21E-01	1.12E-01	9.97E-02	9.63E-01	40	2.91E-12	1.46E-02	4.31E-03	9.95E-01	5	-1.81E-05	3.61E-01	9.96E-01	-6.73E-02	343

Table A.3: Parameter estimates of GARCH(1,1), TGARCH(1,1) and N-GARCH(1,1) models for FTSE100 datasets (Student's t-distribution)

Datasets	GARCH				NA1-GARCH							TGARCH					EGARCH				
	ω	α	β	*df*	ω	α	β	κ	γ	λ	*df*	ω	α^+	α^-	β	*df*	ω	α	β	γ	*df*
SP1	1.31E-06	4.46E-02	9.23E-01	198	1.27E-06	1.07E-01	8.85E-01	8.78E-02	1.13E-01	1.00E+00	46	2.96E-13	1.23E-02	1.28E-02	9.92E-01	13	1.20E-06	8.56E-02	1.00E+00	-3.74E-02	67
SP2	2.24E-06	3.25E-02	9.12E-01	198	1.87E-06	1.28E-01	8.72E-01	1.43E-01	1.99E-01	1.00E+00	62	1.14E-06	5.81E-02	9.59E-08	9.78E-01	6	-1.75E-06	7.25E-02	9.99E-01	-8.33E-02	58
SP3	2.31E-06	3.96E-02	9.10E-01	198	1.03E-06	6.95E-02	8.69E-01	1.34E-01	2.27E-05	9.22E-01	81	3.03E-12	7.67E-02	2.00E-08	9.73E-01	18	1.20E-06	7.85E-02	9.99E-01	-8.22E-02	60
SP4	1.01E-06	5.44E-02	9.35E-01	198	1.91E-06	1.62E-01	8.28E-01	5.33E-02	7.86E-02	9.85E-01	31	4.53E-05	3.60E-02	2.22E-02	9.75E-01	198	-7.11E-07	9.04E-02	9.99E-01	-7.91E-02	52
SP5	9.51E-07	5.57E-02	9.37E-01	198	1.66E-06	2.34E-01	7.56E-01	3.96E-02	1.74E-07	9.40E-01	44	3.98E-08	5.08E-02	5.43E-03	9.82E-01	11	1.50E-06	8.47E-02	9.99E-01	-6.44E-02	65
SP6	3.36E-07	1.32E-02	9.87E-01	198	1.70E-06	1.03E-01	8.28E-01	8.86E-02	1.44E-01	9.17E-01	83	9.42E-08	7.81E-02	3.19E-01	8.72E-01	13	-1.10E-05	2.06E-01	9.97E-01	-7.56E-02	343
SP7	3.96E-06	9.37E-02	9.01E-01	198	1.70E-06	7.21E-02	7.82E-01	8.13E-02	1.67E-04	8.36E-01	50	1.02E-05	1.43E-01	2.45E-01	8.66E-01	197	-2.74E-06	1.61E-01	1.00E+00	-1.16E-01	342
SP8	1.33E-06	8.31E-02	9.17E-01	55	1.62E-06	9.79E-02	8.97E-01	1.78E-01	7.62E-02	9.63E-01	39	4.02E-05	2.49E-02	2.43E-01	9.00E-01	14	-5.47E-06	1.26E-01	1.00E+00	-1.32E-01	343
SP9	6.75E-07	8.29E-02	9.17E-01	198	1.27E-07	9.95E-02	9.00E-01	1.23E-01	1.51E-01	9.69E-01	74	7.29E-14	1.41E-04	2.58E-02	9.90E-01	6	-7.08E-06	1.64E-01	1.00E+00	-1.07E-01	343
SP10	3.18E-06	9.85E-02	8.93E-01	198	1.32E-06	1.11E-01	7.05E-01	1.32E-01	1.28E-01	7.89E-01	50	4.85E-10	8.79E-02	3.50E-01	8.53E-01	13	-1.22E-05	1.69E-01	1.00E+00	-1.17E-01	343
SP11	8.60E-07	1.03E-01	8.97E-01	198	8.90E-07	1.15E-01	6.90E-01	1.29E-01	1.61E-01	7.72E-01	100	8.73E-05	1.50E-01	2.46E-01	8.56E-01	160	-1.34E-05	2.30E-01	9.99E-01	-8.24E-02	343
SP12	1.86E-06	6.98E-02	9.17E-01	198	1.52E-06	8.90E-02	9.10E-01	8.96E-02	9.95E-02	9.92E-01	24	9.20E-06	1.58E-01	2.42E-01	8.59E-01	38	-2.02E-05	1.79E-01	9.99E-01	-5.76E-02	343
SP13	1.95E-06	7.33E-02	9.11E-01	190	1.98E-06	9.64E-02	8.99E-01	5.55E-02	7.55E-02	9.95E-01	59	4.92E-13	2.15E-01	1.57E-02	9.16E-01	9	-7.00E-07	1.59E-01	9.99E-01	-9.94E-02	36
SP14	2.88E-06	8.23E-02	8.91E-01	198	1.52E-06	7.48E-02	8.95E-01	7.92E-02	9.37E-02	9.69E-01	57	4.25E-06	2.50E-01	2.81E-01	8.23E-01	12	-1.02E-05	2.75E-01	9.96E-01	-1.00E-01	343
SP15	2.57E-06	1.33E-01	8.61E-01	198	1.67E-06	1.06E-01	7.10E-01	1.46E-01	8.62E-02	7.84E-01	57	1.78E-06	2.14E-02	3.15E-04	9.94E-01	5	-1.73E-06	2.23E-01	9.98E-01	-8.39E-02	343
SP16	3.33E-06	1.27E-01	8.54E-01	198	1.84E-06	1.26E-01	8.74E-01	1.01E-01	4.36E-02	9.82E-01	51	4.36E-06	5.29E-02	1.54E-04	9.81E-01	26	2.33E-06	2.45E-01	9.98E-01	-7.88E-02	343
SP17	2.80E-06	9.75E-02	8.80E-01	198	2.26E-06	7.52E-02	9.03E-01	1.07E-01	3.23E-02	9.71E-01	69	1.13E-12	1.38E-05	2.96E-02	9.90E-01	10	-5.49E-06	2.03E-01	9.98E-01	-9.40E-02	158
SP18	3.03E-06	1.22E-01	8.55E-01	198	1.49E-06	8.38E-02	8.96E-01	9.35E-02	8.30E-02	9.76E-01	82	2.34E-06	7.82E-04	2.34E-01	9.30E-01	133	4.58E-07	1.82E-01	9.98E-01	-1.06E-01	343
SP19	4.61E-06	1.53E-01	8.13E-01	198	1.90E-06	1.11E-01	8.79E-01	1.14E-01	8.12E-02	9.79E-01	39	2.15E-11	1.78E-02	8.26E-03	9.91E-01	5	-3.11E-06	2.48E-01	9.98E-01	-1.69E-01	343
SP20	5.21E-06	1.65E-01	7.98E-01	198	1.91E-06	1.15E-01	5.58E-01	7.38E-02	1.20E-02	6.59E-01	107	8.76E-11	2.89E-02	1.63E-02	9.82E-01	5	-3.11E-05	2.08E-01	9.98E-01	-1.17E-01	342
SP21	4.40E-06	5.33E-02	8.73E-01	197	2.07E-06	1.12E-01	8.84E-01	6.20E-02	7.77E-02	1.01E+00	68	1.44E-13	2.01E-02	7.47E-03	9.90E-01	6	1.37E-06	-5.32E-02	1.00E+00	-4.19E-02	76
SP22	9.79E-06	9.13E-02	7.47E-01	198	1.67E-06	1.56E-01	8.43E-01	7.46E-02	1.02E-01	9.99E-01	36	2.03E-05	9.03E-02	1.46E-02	9.58E-01	198	1.61E-06	-4.01E-02	9.99E-01	-1.08E-01	29
SP23	1.06E-05	1.29E-01	6.62E-01	198	2.69E-06	3.93E-02	4.35E-01	1.43E-01	8.46E-02	4.92E-01	56	2.55E-06	1.98E-02	8.14E-07	9.93E-01	6	-1.71E-07	1.18E-01	9.99E-01	-1.43E-01	343
SP24	7.63E-06	1.87E-01	6.61E-01	198	1.54E-06	1.23E-01	7.02E-01	8.82E-02	4.42E-02	8.29E-01	65	3.77E-10	2.82E-02	3.17E-04	9.90E-01	5	-8.26E-07	3.25E-01	9.97E-01	-1.69E-01	343

Table A.4: Parameter estimates of GARCH(1,1), TGARCH(1,1) and N-GARCH(1,1) models for SP500 datasets (Student's t-distribution)

A.2 Tables of out-of-sample errors

Datasets	MAE				RMSE			
	GARCH	NA1-GARCH	TGARCH	EGARCH	GARCH	NA1-GARCH	TGARCH	EGARCH
FTSE1	1.19E-03	**1.12E-03**	2.70E-03	1.15E-03	1.79E-03	1.64E-03	3.82E-03	**1.53E-03**
FTSE2	1.90E-03	**1.72E-03**	4.33E-03	1.72E-03	2.60E-03	2.34E-03	5.31E-03	**2.20E-03**
FTSE3	2.48E-03	**1.83E-03**	1.90E-03	2.19E-03	3.17E-03	2.43E-03	**2.40E-03**	2.67E-03
FTSE4	3.39E-03	3.75E-03	**2.87E-03**	3.04E-03	4.74E-03	5.18E-03	**3.98E-03**	4.04E-03
FTSE5	3.06E-03	**2.99E-03**	4.24E-03	3.57E-03	4.19E-03	**4.10E-03**	5.79E-03	4.69E-03
FTSE6	**2.79E-03**	3.03E-03	1.57E-02	3.60E-03	**3.68E-03**	3.97E-03	1.98E-02	4.48E-03
FTSE7	1.90E-03	**1.47E-03**	2.29E-03	2.06E-03	2.55E-03	**2.20E-03**	3.43E-03	3.23E-03
FTSE8	1.78E-03	1.83E-03	**1.34E-03**	1.51E-03	2.03E-03	2.11E-03	**1.63E-03**	1.82E-03
FTSE9	2.12E-03	**1.16E-03**	5.24E-03	1.17E-03	2.43E-03	**1.40E-03**	6.32E-03	1.63E-03
FTSE10	1.84E-03	**1.16E-03**	1.76E-03	1.20E-03	2.28E-03	**1.44E-03**	2.16E-03	1.67E-03
FTSE11	1.21E-03	**1.03E-03**	4.41E-03	1.21E-03	1.43E-03	**1.35E-03**	5.75E-03	1.51E-03
FTSE12	4.46E-03	**1.20E-03**	4.53E-03	2.69E-03	9.46E-03	**1.73E-03**	5.07E-03	5.18E-03
FTSE13	1.95E-03	1.81E-03	**1.71E-03**	2.93E-03	2.91E-03	2.35E-03	**2.20E-03**	4.05E-03
FTSE14	1.86E-03	**1.53E-03**	2.82E-03	1.98E-03	2.47E-03	**2.14E-03**	3.81E-03	2.56E-03
FTSE15	2.37E-03	2.20E-03	1.86E-02	**1.66E-03**	3.29E-03	3.16E-03	2.14E-02	**2.22E-03**
FTSE16	1.08E-03	**9.64E-04**	1.41E-03	1.16E-03	1.34E-03	**1.18E-03**	1.83E-03	1.58E-03
FTSE17	1.63E-03	**1.44E-03**	2.22E-03	1.74E-03	2.69E-03	2.54E-03	2.88E-03	**2.31E-03**
FTSE18	1.32E-03	**8.28E-04**	1.23E-03	1.85E-03	1.59E-03	**1.04E-03**	1.46E-03	2.19E-03
FTSE19	1.19E-03	**1.19E-03**	3.36E-03	2.41E-02	1.45E-03	**1.37E-03**	3.59E-03	3.82E-02
FTSE20	**1.14E-03**	1.21E-03	1.73E-03	2.18E-03	**1.33E-03**	1.65E-03	2.37E-03	4.28E-03
FTSE21	1.99E-03	**1.15E-03**	2.35E-03	1.23E-03	2.56E-03	**1.64E-03**	2.67E-03	1.66E-03
FTSE22	1.55E-03	**1.19E-03**	1.39E-03	1.22E-03	1.80E-03	**1.39E-03**	1.63E-03	1.44E-03
FTSE23	1.51E-03	**1.06E-03**	2.01E-03	1.44E-03	1.84E-03	**1.34E-03**	2.66E-03	1.80E-03
FTSE24	1.75E-03	1.45E-03	6.75E-03	**1.35E-03**	2.50E-03	2.16E-03	1.03E-02	**1.76E-03**

Table A.5: MAE and RMSE for GARCH, NA1-GARCH, TGARCH and EGARCH models for FTSE100 datasets (Normal distribution)

Datasets	MAE				RMSE			
	GARCH	**NA1-GARCH**	**TGARCH**	**EGARCH**	**GARCH**	**NA1-GARCH**	**TGARCH**	**EGARCH**
SP1	1.89E-03	1.77E-03	**1.26E-03**	7.61E-03	2.34E-03	2.25E-03	**1.68E-03**	8.82E-03
SP2	2.94E-03	**1.62E-03**	3.12E-02	7.62E-03	3.45E-03	**2.05E-03**	4.73E-02	8.82E-03
SP3	3.16E-03	**3.04E-03**	5.82E-03	**3.00E-03**	3.60E-03	3.53E-03	6.67E-03	3.56E-03
SP4	**3.24E-03**	3.61E-03	4.92E-03	1.21E+00	**4.35E-03**	5.33E-03	6.79E-03	2.15E+00
SP5	**2.88E-03**	3.36E-03	1.19E-02	5.87E-03	**4.10E-03**	4.84E-03	1.46E-02	7.98E-03
SP6	1.08E-02	**4.08E-03**	1.54E-02	8.87E-03	1.31E-02	**5.64E-03**	1.95E-02	9.95E-03
SP7	2.09E-03	**1.75E-03**	1.04E-02	1.82E-03	2.90E-03	**2.55E-03**	1.27E-02	2.85E-03
SP8	1.73E-03	2.34E-03	5.17E-03	**1.49E-03**	2.26E-03	2.89E-03	6.37E-03	**1.91E-03**
SP9	1.88E-03	1.33E-03	1.66E-03	**1.30E-03**	2.18E-03	1.73E-03	2.07E-03	**1.69E-03**
SP10	1.99E-03	1.80E-03	3.91E-03	**1.46E-03**	2.61E-03	2.26E-03	5.33E-03	**1.86E-03**
SP11	1.35E-03	**1.02E-03**	1.34E-03	1.18E-03	1.59E-03	**1.27E-03**	2.03E-03	1.44E-03
SP12	4.97E-03	**2.33E-03**	4.68E-03	2.66E-03	1.04E-02	**4.52E-03**	6.60E-03	5.59E-03
SP13	2.68E-03	**1.93E-03**	6.35E-03	2.01E-03	5.23E-03	**2.63E-03**	7.91E-03	2.66E-03
SP14	**1.71E-03**	1.76E-03	7.36E-03	3.80E-03	**2.40E-03**	2.45E-03	8.91E-03	5.44E-03
SP15	1.85E-03	**1.77E-03**	2.14E-03	2.19E-03	**3.52E-03**	3.77E-03	**3.21E-03**	3.92E-03
SP16	1.38E-03	**1.09E-03**	1.66E-03	1.34E-03	1.93E-03	**1.60E-03**	2.22E-03	1.79E-03
SP17	**1.65E-03**	1.68E-03	1.66E-03	2.55E+01	2.82E-03	3.02E-03	**2.24E-03**	7.71E+01
SP18	1.31E-03	**1.19E-03**	3.03E-03	1.32E-03	1.53E-03	**1.37E-03**	3.59E-03	1.63E-03
SP19	1.37E-03	1.31E-03	3.01E-03	**1.26E-03**	1.59E-03	**1.53E-03**	3.88E-03	1.61E-03
SP20	**1.33E-03**	1.91E-03	2.35E-03	6.34E-01	**1.53E-03**	2.19E-03	3.07E-03	6.55E-01
SP21	2.15E-03	1.60E-03	1.60E-03	**1.49E-03**	3.46E-03	2.50E-03	**2.05E-03**	2.31E-03
SP22	1.64E-03	1.50E-03	8.56E-03	**1.03E-03**	1.88E-03	1.79E-03	1.17E-02	**1.24E-03**
SP23	1.48E-03	**9.34E-04**	3.24E-03	3.50E-02	1.77E-03	**1.11E-03**	3.86E-03	5.62E-02
SP24	1.87E-03	**1.79E-03**	2.16E-03	1.84E-03	2.77E-03	2.64E-03	3.03E-03	**2.50E-03**

Table A.6: MAE and RMSE for GARCH, NA1-GARCH, TGARCH and EGARCH models for SP500 datasets (Normal distribution)

Datasets	MAE				RMSE			
	GARCH	NA1-GARCH	TGARCH	EGARCH	GARCH	NA1-GARCH	TGARCH	EGARCH
FTSE1	1.20E-03	**1.00E-03**	3.05E-03	1.15E-03	1.80E-03	**1.29E-03**	3.73E-03	1.52E-03
FTSE2	1.76E-03	**1.47E-03**	3.50E-03	1.72E-03	2.39E-03	**1.98E-03**	4.35E-03	2.23E-03
FTSE3	2.48E-03	**2.07E-03**	4.96E-03	2.18E-03	3.17E-03	2.69E-03	6.36E-03	**2.66E-03**
FTSE4	3.40E-03	**2.95E-03**	8.93E-03	3.13E-03	4.75E-03	**4.15E-03**	1.21E-02	4.15E-03
FTSE5	3.06E-03	**2.53E-03**	8.19E-03	3.55E-03	4.20E-03	**3.53E-03**	1.07E-02	4.69E-03
FTSE6	2.78E-03	**2.72E-03**	4.48E-03	3.38E-03	3.68E-03	3.72E-03	6.43E-03	**3.28E-03**
FTSE7	1.39E-03	**1.33E-03**	8.40E-03	2.01E-03	2.14E-03	**2.00E-03**	9.53E-03	3.15E-03
FTSE8	2.34E-03	**1.14E-03**	4.77E-03	1.50E-03	2.64E-03	**1.51E-03**	5.98E-03	1.81E-03
FTSE9	1.89E-03	1.67E-03	4.72E-03	**1.16E-03**	2.17E-03	2.01E-03	6.01E-03	**1.60E-03**
FTSE10	1.82E-03	1.32E-03	1.77E-03	**1.08E-03**	2.27E-03	1.78E-03	2.21E-03	**1.50E-03**
FTSE11	1.11E-03	**1.02E-03**	2.91E-03	1.15E-03	1.52E-03	**1.38E-03**	4.47E-03	1.45E-03
FTSE12	4.44E-03	**1.13E-03**	4.14E-03	2.69E-03	9.34E-03	**1.62E-03**	6.39E-03	5.42E-03
FTSE13	1.95E-03	**1.86E-03**	3.08E-03	1.87E-03	2.91E-03	2.53E-03	4.00E-03	**2.35E-03**
FTSE14	**1.36E-03**	1.92E-03	3.72E-03	1.95E-03	**1.82E-03**	2.75E-03	5.40E-03	2.59E-03
FTSE15	2.38E-03	**1.83E-03**	4.99E-03	3.58E-01	3.31E-03	**2.55E-03**	7.16E-03	1.67E+00
FTSE16	1.07E-03	**8.77E-04**	2.39E-03	1.20E-03	1.33E-03	**1.13E-03**	3.06E-03	1.61E-03
FTSE17	1.61E-03	**1.26E-03**	1.64E-03	1.72E-03	2.66E-03	**2.05E-03**	2.16E-03	2.27E-03
FTSE18	1.32E-03	**1.22E-03**	2.92E-03	1.83E-03	1.59E-03	**1.51E-03**	3.52E-03	2.17E-03
FTSE19	1.19E-03	**9.29E-04**	2.65E-03	6.77E-03	1.44E-03	**1.12E-03**	3.40E-03	9.90E-03
FTSE20	**1.13E-03**	1.13E-03	2.47E-03	1.43E-03	**1.31E-03**	1.71E-03	3.70E-03	2.17E-03
FTSE21	1.97E-03	**9.55E-04**	1.00E-03	1.40E-03	2.53E-03	1.36E-03	**1.22E-03**	2.01E-03
FTSE22	1.46E-03	1.23E-03	2.53E-03	**1.08E-03**	1.68E-03	1.46E-03	3.46E-03	**1.26E-03**
FTSE23	1.45E-03	**1.30E-03**	3.04E-03	1.38E-03	1.75E-03	**1.63E-03**	3.76E-03	1.72E-03
FTSE24	1.76E-03	**1.69E-03**	4.56E-03	2.34E-03	2.51E-03	**2.40E-03**	5.71E-03	3.26E-03

Table A.7: MAE and RMSE for GARCH, NA1-GARCH, TGARCH and EGARCH models for FTSE100 datasets (Student's t-distribution)

Datasets	MAE				RMSE			
	GARCH	NA1-GARCH	TGARCH	EGARCH	GARCH	NA1-GARCH	TGARCH	EGARCH
SP1	1.70E-03	**1.49E-03**	3.41E-03	1.94E-03	2.10E-03	**2.05E-03**	4.04E-03	2.31E-03
SP2	2.87E-03	**1.43E-03**	2.67E-03	1.81E-03	3.37E-03	**1.82E-03**	3.24E-03	2.30E-03
SP3	3.06E-03	**1.22E-03**	3.00E-03	2.86E-03	3.49E-03	**1.56E-03**	4.00E-03	3.41E-03
SP4	3.12E-03	**2.97E-03**	6.36E-03	4.50E-03	4.19E-03	**4.17E-03**	8.93E-03	6.37E-03
SP5	**2.82E-03**	4.38E-03	8.13E-03	5.33E-03	**4.03E-03**	5.99E-03	1.05E-02	7.26E-03
SP6	9.89E-03	**2.73E-03**	7.85E-03	4.89E-03	1.14E-02	**3.91E-03**	1.12E-02	6.09E-03
SP7	2.08E-03	**1.55E-03**	2.12E-03	1.85E-03	2.90E-03	**2.30E-03**	3.39E-03	2.86E-03
SP8	1.43E-03	**1.40E-03**	2.76E-03	1.59E-03	1.99E-03	**1.89E-03**	3.77E-03	2.05E-03
SP9	1.04E-03	**9.67E-04**	4.31E-03	1.34E-03	1.27E-03	**1.20E-03**	5.48E-03	1.73E-03
SP10	1.98E-03	**1.36E-03**	2.99E-03	1.52E-03	2.54E-03	**1.78E-03**	4.08E-03	1.95E-03
SP11	**8.95E-04**	1.20E-03	1.76E-03	1.21E-03	**1.11E-03**	1.51E-03	2.23E-03	1.51E-03
SP12	4.44E-03	3.51E-03	**1.95E-03**	2.91E-03	9.29E-03	7.67E-03	**2.90E-03**	6.40E-03
SP13	2.66E-03	**1.40E-03**	2.76E-03	2.15E-03	5.15E-03	**1.93E-03**	3.84E-03	2.99E-03
SP14	1.71E-03	**1.32E-03**	2.93E-03	2.76E-03	2.40E-03	**1.98E-03**	4.37E-03	3.83E-03
SP15	**1.77E-03**	1.85E-03	6.26E-03	2.14E-03	**3.60E-03**	3.83E-03	8.79E-03	3.82E-03
SP16	1.36E-03	**1.22E-03**	2.09E-03	1.35E-03	1.90E-03	**1.69E-03**	2.55E-03	1.79E-03
SP17	1.64E-03	**1.33E-03**	2.43E-03	1.46E-03	2.83E-03	2.22E-03	3.18E-03	**1.99E-03**
SP18	1.17E-03	**9.67E-04**	1.82E-03	1.31E-03	1.38E-03	**1.15E-03**	2.39E-03	1.60E-03
SP19	1.34E-03	**1.23E-03**	2.53E-03	1.62E-03	1.56E-03	**1.44E-03**	3.34E-03	2.14E-03
SP20	1.44E-03	1.34E-03	1.84E-03	**1.18E-03**	1.66E-03	1.58E-03	2.74E-03	**1.50E-03**
SP21	2.10E-03	**1.25E-03**	2.62E-03	2.09E-01	3.40E-03	**1.73E-03**	3.14E-03	3.41E-01
SP22	1.66E-03	**1.06E-03**	1.57E-03	1.67E-03	1.90E-03	**1.32E-03**	1.89E-03	2.06E-03
SP23	1.46E-03	**1.40E-03**	2.94E-03	1.42E-03	1.75E-03	**1.61E-03**	3.66E-03	1.65E-03
SP24	1.83E-03	**1.21E-03**	3.38E-03	2.01E-03	2.69E-03	**1.74E-03**	4.70E-03	2.91E-03

Table A.8: MAE and RMSE for GARCH, NA1-GARCH, TGARCH and EGARCH models for SP00 datasets (Student's t-distribution)

A.3 Tables of in-sample model comparison

Datasets	LLH				AIC				BIC			
	GARCH	**NA1-GARCH**	**TGARCH**	**EGARCH**	**GARCH**	**NA1-GARCH**	**TGARCH**	**EGARCH**	**GARCH**	**NA1-GARCH**	**TGARCH**	**EGARCH**
FTSE1	2283.17	1823.10	1802.88	*2284.63*	-4560.33	-3634.20	-3597.77	**-4561.269**	**-4547.69**	-3608.91	-3580.91	-4544.41
FTSE2	2263.31	1802.73	6474.74	*2264.05*	-4520.61	-3593.46	-2957.48	-4520.099	**-4507.97**	-3568.18	-2974.34	-4503.24
FTSE3	*2243.22*	1780.75	1786.27	2242.46	-4480.43	-3549.51	-3564.53	-4476.923	**-4467.79**	-3524.22	-3547.67	-4460.06
FTSE4	*2136.85*	1675.93	1686.50	2136.16	-4267.70	-3339.85	-3365.01	-4264.317	**-4255.06**	-3314.56	-3348.15	-4247.46
FTSE5	2084.64	1624.45	1642.54	*2089.18*	-4163.27	-3236.91	-3277.07	**-4170.361**	-4150.63	-3211.62	-3260.21	**-4153.50**
FTSE6	2042.09	1581.88	461.49	*2048.17*	-4078.18	-3151.76	-930.99	**-4088.347**	-4065.53	-3126.48	-947.85	**-4071.49**
FTSE7	1811.69	1349.76	1364.87	*1820.20*	-3617.39	-2687.53	-2721.74	**-3632.397**	-3604.74	-2662.24	-2704.88	**-3615.54**
FTSE8	1774.22	1317.09	1327.07	*1785.73*	-3542.45	-2622.19	-2646.13	**-3563.451**	-3529.81	-2596.90	-2629.27	**-3546.59**
FTSE9	1793.33	1332.60	244.34	*1801.44*	-3580.67	-2653.19	-480.69	**-3594.884**	-3568.02	-2627.91	-463.83	**-3578.03**
FTSE10	1824.94	1363.61	1379.83	*1832.48*	-3643.87	-2715.22	-2751.65	**-3656.965**	-3631.23	-2689.93	-2734.80	**-3640.11**
FTSE11	1841.35	1379.07	1308.16	*1852.04*	-3676.69	-2746.14	-2608.32	**-3696.086**	-3664.05	-2720.85	-2591.46	**-3679.23**
FTSE12	1869.65	1400.62	1225.36	*1873.82*	-3733.30	-2789.23	2458.71	**-3739.632**	-3720.66	-2763.95	-2475.57	**-3722.77**
FTSE13	*1955.65*	1493.85	1502.44	1922.65	**-3905.30**	-2975.70	-2996.88	-3837.302	**-3892.66**	-2950.41	-2980.02	-3820.44
FTSE14	2038.52	1576.59	1598.08	*2045.45*	-4071.03	-3141.19	-3188.17	**-4082.898**	-4058.39	-3115.90	-3171.31	**-4066.04**
FTSE15	2056.31	1595.14	1550.11	*2061.70*	-4106.63	-3178.29	-3138.23	**-4115.405**	-4093.98	-3153.00	-3105.09	**-4098.55**
FTSE16	1998.28	1537.81	1558.38	*2007.38*	-3990.57	-3063.62	-3108.75	**-4006.762**	-3977.92	-3038.34	-3091.89	**-3989.90**
FTSE17	2002.93	1541.38	1414.89	*2013.19*	-3999.87	-3070.75	-2821.77	**-4018.38**	-3987.22	-3045.47	-2804.91	**-4001.52**
FTSE18	2027.46	1565.94	1583.40	*2041.60*	-4048.92	-3119.87	-3158.81	**-4075.209**	-4036.27	-3094.58	-3141.95	**-4058.35**
FTSE19	2047.22	1586.70	1594.89	*2072.73*	-4088.43	-3161.40	-3197.78	**-4137.452**	-4075.79	-3136.11	-3214.63	**-4120.59**
FTSE20	*2061.09*	1599.93	1605.40	2053.26	**-4116.17**	-3187.85	-3202.81	-4098.517	**-4103.53**	-3162.57	-3185.95	-4081.66
FTSE21	*2131.20*	1666.48	1667.64	2128.70	**-4256.40**	-3320.95	-3327.28	-4249.405	**-4243.76**	-3295.66	-3310.43	-4232.55
FTSE22	*2172.11*	1704.93	1711.22	2170.11	**-4338.22**	-3397.85	-3414.45	-4332.213	**-4325.58**	-3372.56	-3397.59	-4315.35
FTSE23	*2223.17*	1757.71	1772.24	2219.44	**-4440.34**	-3503.42	-3536.48	-4430.889	**-4427.69**	-3478.13	-3519.62	-4414.03
FTSE24	*2253.07*	1791.13	1677.14	2241.86	**-4500.15**	-3570.27	-3346.28	-4475.717	**-4487.50**	-3544.98	-3329.42	-4458.86

Table A.9: Likelihood, AIC and BIC for GARCH, NA1-GARCH, TGARCH and EGARCH models for FTSE100 datasets (Normal distribution)

Datasets	LLH				AIC				BIC			
	GARCH	NA1-GARCH	TGARCH	EGARCH	GARCH	NA1-GARCH	TGARCH	EGARCH	GARCH	NA1-GARCH	TGARCH	EGARCH
SP1	2283.22	1820.97	1803.27	*2301.15*	-4560.43	-3629.93	-3598.53	**-4594.30**	-4547.79	-3604.64	-3581.67	**-4577.44**
SP2	2270.57	1803.97	1308.68	*2280.65*	-4535.13	-3595.93	-2609.36	**-4553.31**	-4522.49	-3570.65	-2592.50	**-4536.45**
SP3	*2257.58*	1796.26	60.37	2254.59	**-4509.16**	-3580.51	-112.74	-4501.19	**-4496.52**	-3555.22	-95.89	-4484.33
SP4	2161.71	1697.78	1705.76	*2183.69*	-4317.42	-3383.56	-3403.52	**-4359.38**	-4304.77	-3358.28	-3386.66	**-4342.52**
SP5	2117.57	1655.32	1598.69	*2118.54*	**-4229.14**	-3298.64	-3189.38	-4229.08	**-4216.49**	-3273.35	-3172.52	-4212.23
SP6	*2068.93*	1600.98	1269.89	2065.67	**-4131.86**	-3189.97	-3547.79	-4123.35	**-4119.22**	-3164.68	-3564.65	-4106.49
SP7	1791.58	1333.35	1187.62	*1792.25*	**-3577.16**	-2654.71	-2383.23	-3576.51	**-3564.52**	-2629.42	-2400.09	-3559.65
SP8	1753.07	1301.21	1679.31	*1760.91*	-3500.14	-2590.41	-3466.63	**-3513.82**	-3487.49	-2565.13	-3383.49	**-3496.96**
SP9	1766.74	1305.42	1318.09	*1772.26*	-3527.48	-2598.84	-2628.18	**-3536.52**	-3514.83	-2573.56	-2611.32	**-3519.66**
SP10	1787.32	1326.33	1654.53	*1793.74*	-3568.63	-2640.66	-3317.06	**-3579.47**	-3555.99	-2615.37	-3033.92	**-3562.62**
SP11	1796.19	1337.88	1335.61	*1798.33*	-3586.38	-2663.76	-2663.23	**-3588.66**	**-3573.73**	-2638.47	-2646.37	-3571.80
SP12	*1829.27*	1365.68	824.52	1828.36	**-3652.53**	-2719.35	-1641.03	-3648.72	**-3639.89**	-2694.06	-1624.17	-3631.87
SP13	*1935.71*	1470.11	1764.31	1934.26	**-3865.41**	-2928.22	-1536.63	-3860.52	**-3852.77**	-2902.93	-3553.48	-3843.66
SP14	*2052.63*	1587.09	1513.80	1984.80	**-4099.26**	-3162.19	-2335.60	-3961.59	**-4086.61**	-3136.90	-3052.46	-3944.73
SP15	*2035.92*	1576.65	1590.62	2035.51	**-4065.84**	-3141.30	-3173.23	-4063.02	**-4053.20**	-3116.01	-3156.38	-4046.16
SP16	*1992.05*	1537.30	1543.82	1989.18	**-3978.09**	-3062.60	-3079.65	-3970.35	**-3965.45**	-3037.31	-3062.79	-3953.49
SP17	*2002.87*	1546.57	1318.24	1827.72	**-3999.74**	-3081.14	-2628.49	-3647.45	**-3987.10**	-3055.86	-2611.63	-3630.59
SP18	2052.38	1594.81	1188.89	*2053.51*	-4098.77	-3177.61	-2369.79	**-4099.01**	**-4086.12**	-3152.33	-2352.93	-4082.16
SP19	2052.28	1581.92	1559.30	*2058.31*	-4098.56	-3151.85	-3110.60	**-4108.62**	-4085.92	-3126.56	-3093.75	**-4091.76**
SP20	*2059.73*	1590.46	1604.09	2035.51	**-4113.46**	-3168.92	-3200.17	-4063.02	**-4100.82**	-3143.64	-3183.31	-4046.16
SP21	2156.70	1695.88	1576.50	*2160.46*	-4307.41	-3379.76	-3144.99	**-4312.92**	-4294.76	-3354.47	-3128.13	**-4296.06**
SP22	2186.52	1723.17	1438.77	*2192.10*	-4367.05	-3434.35	-2869.55	**-4376.19**	-4354.40	-3409.06	-2852.69	**-4359.34**
SP23	2230.86	1764.25	691.07	*2240.76*	-4455.72	-3516.50	-1374.13	**-4473.52**	-4443.07	-3491.22	-1357.27	**-4456.66**
SP24	*2252.67*	1792.72	1812.10	2247.72	**-4499.34**	-3573.44	-3616.20	-4487.43	**-4486.69**	-3548.15	-3599.34	-4470.57

Table A.10: Likelihood, AIC and BIC for GARCH, NA1-GARCH, TGARCH and EGARCH models for SP500 datasets (Normal distribution)

Datasets	LLH				AIC				BIC			
	GARCH	**NA1-GARCH**	**TGARCH**	**EGARCH**	**GARCH**	**NA1-GARCH**	**TGARCH**	**EGARCH**	**GARCH**	**NA1-GARCH**	**TGARCH**	**EGARCH**
FTSE1	2280.48	2248.41	2015.79	*2284.63*	-4554.97	-4484.82	-4023.58	**-4561.27**	-4542.32	-4459.54	-4006.72	**-4544.41**
FTSE2	2260.69	2234.12	2010.47	*2264.05*	-4515.39	-4456.24	-4012.94	**-4520.10**	-4502.74	-4430.95	-3996.09	**-4503.24**
FTSE3	*2240.77*	2236.12	2013.35	2237.82	**-4475.55**	-4460.24	-4018.71	-4467.63	**-4462.90**	-4434.96	-4001.85	-4450.77
FTSE4	2134.33	2129.54	1849.10	*2136.16*	-4262.66	-4247.09	-3690.20	**-4264.32**	**-4250.02**	-4221.80	-3673.35	-4247.46
FTSE5	2082.50	2078.14	1779.49	*2089.18*	-4159.01	-4144.28	-3550.98	**-4170.36**	-4146.36	-4119.00	-3534.12	**-4153.50**
FTSE6	*2040.00*	2036.41	1957.49	1961.92	**-4074.00**	-4060.81	-3906.99	-3915.86	**-4061.36**	-4035.52	-3890.13	-3899.00
FTSE7	1806.38	1803.81	1570.88	*1820.20*	-3606.76	-3595.61	-3133.76	**-3632.40**	-3594.12	-3570.33	-3116.90	**-3615.54**
FTSE8	1775.50	1769.02	1509.89	*1782.40*	-3545.00	-3526.04	-3011.77	**-3556.79**	-3532.35	-3500.76	-2994.91	**-3539.93**
FTSE9	1790.50	1788.17	1466.14	*1801.44*	-3575.00	-3564.34	-2924.28	**-3594.88**	-3562.36	-3539.05	-2907.43	**-3578.03**
FTSE10	1822.54	1808.89	1792.75	*1832.48*	-3639.09	-3605.79	-3577.50	**-3656.97**	-3626.44	-3580.50	-3560.65	**-3640.11**
FTSE11	1837.89	1828.19	1511.30	*1852.04*	-3669.78	-3644.38	-3014.61	**-3696.09**	-3657.14	-3619.09	-2997.75	**-3679.23**
FTSE12	1867.24	1848.97	1804.66	*1873.82*	-3728.47	-3685.94	-3601.33	**-3739.63**	-3715.83	-3660.65	-3584.47	**-3722.77**
FTSE13	*1953.18*	1948.75	1785.67	1922.65	**-3900.35**	-3885.49	-3563.34	-3837.30	**-3887.71**	-3860.20	-3546.48	-3820.44
FTSE14	2034.71	2030.11	2020.47	*2045.45*	-4063.42	-4048.21	-4032.94	**-4082.90**	-4050.77	-4022.92	-4016.08	**-4066.04**
FTSE15	2053.74	2042.89	1972.93	*2061.70*	-4101.48	-4073.78	-3937.85	**-4115.41**	-4088.84	-4048.49	-3921.00	**-4098.55**
FTSE16	1995.69	1988.19	1777.83	*2005.30*	-3985.39	-3964.38	-3547.66	**-4002.61**	-3972.74	-3939.10	-3530.80	**-3985.75**
FTSE17	2000.30	1997.39	1988.29	*2013.19*	-3994.59	-3982.79	-3968.57	**-4018.38**	-3981.95	-3957.50	-3951.71	**-4001.52**
FTSE18	2024.74	2020.30	1684.85	*2041.60*	-4043.48	-4028.59	-3361.69	**-4075.21**	-4030.84	-4003.30	-3344.83	**-4058.35**
FTSE19	2044.70	2035.70	1750.34	*2070.09*	-4083.40	-4059.40	-3492.69	**-4132.18**	-4070.76	-4034.11	-3475.83	**-4115.32**
FTSE20	*2058.60*	2052.37	1789.58	2053.26	**-4111.20**	-4092.74	-3571.15	-4098.52	**-4098.55**	-4067.46	-3554.29	-4081.66
FTSE21	*2128.85*	2120.57	2091.36	2126.24	**-4251.70**	-4229.14	-4174.71	-4244.48	**-4239.06**	-4203.85	-4157.86	-4227.62
FTSE22	2169.79	2162.60	2084.84	*2170.11*	**-4333.57**	-4313.20	-4161.68	-4332.21	**-4320.93**	-4287.91	-4144.82	-4315.35
FTSE23	*2220.61*	2215.61	1963.18	2219.44	**-4435.23**	-4419.23	-3918.36	-4430.89	**-4422.58**	-4393.94	-3901.50	-4414.03
FTSE24	*2251.12*	2239.60	1841.49	2241.86	**-4496.25**	-4467.20	-3674.98	-4475.72	**-4483.61**	-4441.91	-3658.12	-4458.86

Table A.11: Likelihood, AIC and BIC for GARCH, NA1-GARCH, TGARCH and EGARCH models for FTSE100 datasets (Student's t-distribution)

Datasets	LLH				AIC				BIC			
	GARCH	**NA1-GARCH**	**TGARCH**	**EGARCH**	**GARCH**	**NA1-GARCH**	**TGARCH**	**EGARCH**	**GARCH**	**NA1-GARCH**	**TGARCH**	**EGARCH**
SP1	*2280.31*	2268.77	2133.16	2116.52	**-4554.63**	-4525.54	-4258.32	-4225.06	**-4541.99**	-4500.26	-4241.46	-4208.20
SP2	*2269.48*	2257.42	2041.86	2266.51	**-4532.97**	-4502.83	-4075.71	-4525.03	**-4520.33**	-4477.55	-4058.85	-4508.17
SP3	2256.46	2249.12	2120.18	*2256.56*	**-4506.93**	-4486.25	-4232.36	-4505.12	**-4494.28**	-4460.96	-4215.50	-4488.26
SP4	2160.79	2148.14	2152.87	*2164.16*	-4315.59	-4284.28	-4297.73	**-4320.32**	-4302.94	-4258.99	-4280.88	**-4303.46**
SP5	2116.63	2097.02	2016.96	*2119.81*	-4227.27	-4182.05	-4025.92	**-4231.61**	-4214.63	-4156.76	-4009.06	**-4214.75**
SP6	*2067.14*	2049.44	1999.18	2065.67	**-4128.29**	-4086.88	-3990.37	-4123.35	**-4115.64**	-4061.59	-3973.51	-4106.49
SP7	1789.06	1783.83	1767.43	*1792.25*	**-3572.13**	-3555.66	-3526.85	**-3576.51**	-3559.48	-3530.37	-3510.00	**-3559.65**
SP8	1744.29	1748.21	1673.15	*1760.91*	-3482.58	-3484.42	-3338.29	**-3513.82**	-3469.94	-3459.13	-3321.43	**-3496.96**
SP9	1761.01	1757.09	1446.05	*1772.26*	-3516.02	-3502.19	-2884.10	**-3536.52**	-3503.38	-3476.90	-2867.24	**-3519.66**
SP10	1784.79	1766.95	1701.76	*1793.74*	-3563.57	-3521.90	-3395.52	**-3579.47**	-3550.93	-3496.61	-3378.66	**-3562.62**
SP11	1791.54	1786.76	1784.05	*1798.33*	-3577.07	-3561.51	-3560.11	**-3588.66**	-3564.43	-3536.23	-3543.25	**-3571.80**
SP12	1826.34	1812.22	1799.22	*1828.36*	-3646.67	-3612.44	-3590.44	**-3648.72**	**-3634.03**	-3587.16	-3573.58	-3631.87
SP13	*1933.20*	1920.32	1785.28	1930.10	**-3860.41**	-3828.64	-3562.57	-3852.21	**-3847.76**	-3803.35	-3545.71	-3835.35
SP14	*2050.30*	2034.34	1965.49	1984.80	**-4094.60**	-4056.67	-3922.98	-3961.59	**-4081.96**	-4031.38	-3906.12	-3944.73
SP15	2033.15	2026.85	1690.19	*2035.51*	-4060.30	-4041.69	-3372.37	**-4063.02**	**-4047.65**	-4016.41	-3355.51	-4046.16
SP16	*1989.93*	1985.42	1915.04	1989.18	**-3973.86**	-3958.85	-3822.08	-3970.35	**-3961.21**	-3933.56	-3805.22	-3953.49
SP17	*2000.76*	1998.17	1857.69	2000.31	**-3995.52**	-3984.35	-3707.38	-3647.45	**-3982.87**	-3959.06	-3690.52	-3975.76
SP18	2049.80	2048.31	2025.48	*2053.51*	-4093.60	-4084.62	-4042.95	**-4099.01**	-4080.95	-4059.33	-4026.09	**-4082.16**
SP19	2049.50	2030.01	1690.33	*2058.31*	-4093.00	-4048.02	-3372.66	**-4108.62**	-4080.36	-4022.73	-3355.80	**-4091.76**
SP20	*2055.74*	2049.89	1778.69	2035.51	**-4105.49**	-4087.78	-3549.37	-4063.02	**-4092.84**	-4062.49	-3532.52	-4046.16
SP21	2154.37	2147.57	1940.28	*2159.48*	-4302.74	-4283.13	-3872.57	**-4310.96**	-4290.10	-4257.84	-3855.71	**-4294.10**
SP22	*2183.80*	2165.67	2150.52	2161.27	**-4361.60**	-4319.34	-4293.05	-4314.53	**-4348.96**	-4294.05	-4276.19	-4297.68
SP23	2228.53	2221.85	1962.72	*2240.76*	-4451.06	-4431.69	-3917.44	**-4473.52**	-4438.42	-4406.41	-3900.59	**-4456.66**
SP24	*2248.40*	2239.04	1927.96	2247.72	**-4490.81**	-4466.09	-3847.91	-4487.43	**-4478.16**	-4440.80	-3831.06	-4470.57

Table A.12: Likelihood, AIC and BIC for GARCH, NA1-GARCH, TGARCH and EGARCH models for SP500 datasets (Student's t-distribution)

A.4 Tables of VaR Backtesting

Note: In the tables from A.13 to A.20, the entry "Accepted" stands for the case where the model cannot be rejected by the corresponding hypothesis test.

Risk Model	Confidence Level	T	N	Kupiec test	Kupiec Outcome	Christ. test	Christ. Outcome
GARCH	95%	250	13	2.08e-2	Accepted	1.43	Accepted
	99%	250	2	1.08e-1	Accepted	3.24e-2	Accepted
TGARCH	95%	250	17	1.54	Accepted	5.92e-1	Accepted
	99%	250	5	1.96	Accepted	2.05e-1	Accepted
EGARCH	95%	250	19	3.09	Accepted	1.56	Accepted
	99%	250	12	1.90e1	**Rejected**	2.84e-1	Accepted
NA1-GARCH	95%	250	17	1.54	Accepted	5.92e-1	Accepted
	99%	250	4	7.69e-1	Accepted	1.31e-1	Accepted

Table A.13: Summary of the Backtesting data for dataset (FTSE18) From 31 August 2012 To 15 August 2013, using four different models when the distribution is assumed to be normal. (T = total number of observations; N = number of exceptions).

Risk Model	Confidence Level	T	N	Kupiec test	Kupiec Outcome	Christ. test	Christ. Outcome
GARCH	95%	250	11	1.97e-1	Accepted	1.02	Accepted
	99%	250	3	9.49e-2	Accepted	7.32e-2	Accepted
TGARCH	95%	250	15	4.96e-1	Accepted	4.49	**Rejected**
	99%	250	11	1.59e1	**Rejected**	3.85	**Rejected**
EGARCH	95%	250	11	1.97e-1	Accepted	1.02	Accepted
	99%	250	5	1.96	Accepted	2.05e-1	Accepted
NA1-GARCH	95%	250	12	2.13e-2	Accepted	1.22	Accepted
	99%	250	5	1.96	Accepted	2.05e-1	Accepted

Table A.14: Summary of the Backtesting data for dataset (SP18) From 25 September 2012 To 24 September 2013, using four different models when the distribution is assumed to be normal. (T = total number of observations; N = number of exceptions).

Risk Model	Confidence Level	T	N	Kupiec test	Kupiec Outcome	Christ. test	Christ. Outcome
GARCH	95%	250	13	2.08e-2	Accepted	1.43	Accepted
	99%	250	2	1.08e-1	Accepted	3.24e-2	Accepted
TGARCH	95%	250	36	3.16e1	**Rejected**	3.32	Accepted
	99%	250	26	7.71e1	**Rejected**	6.83e-1	Accepted
EGARCH	95%	250	19	3.09	Accepted	1.56	Accepted
	99%	250	12	1.90e1	**Rejected**	2.50	Accepted
NA1-GARCH	95%	250	13	2.08e-2	Accepted	1.43	Accepted
	99%	250	2	1.08e-1	Accepted	3.24e-2	Accepted

Table A.15: Summary of the Backtesting data for dataset (FTSE18) From 31 August 2012 To 15 August 2013, using four different models when the distribution is assumed to be student's t-distribution. (T = total number of observations; N = number of exceptions).

Risk Model	Confidence Level	T	N	Kupiec test	Kupiec Outcome	Christ. test	Christ. Outcome
GARCH	95%	250	12	2.13e-2	Accepted	1.22	Accepted
	99%	250	6	3.56	Accepted	2.96e-1	Accepted
TGARCH	95%	250	13	2.08e-2	Accepted	1.43	Accepted
	99%	250	6	3.56	Accepted	2.96e-1	Accepted
EGARCH	95%	250	11	1.97e-1	Accepted	1.02	Accepted
	99%	250	5	1.96	Accepted	2.05e-1	Accepted
NA1-GARCH	95%	250	13	2.08e-2	Accepted	1.43	Accepted
	99%	250	6	3.56	Accepted	2.96e-1	Accepted

Table A.16: Summary of the Backtesting data for dataset (SP18) From 25 September 2012 To 24 September 2013, using four different models when the distribution is assumed to be student's t-distribution. (T = total number of observations; N = number of exceptions).

Risk Model	Confidence Level	T	N	Kupiec test	Kupiec Outcome	Christ. test	Christ. Outcome
GARCH	95%	250	12	2.13e-2	Accepted	2.84e-1	Accepted
	99%	250	4	7.69e-1	Accepted	1.31e-1	Accepted
TGARCH	95%	250	11	1.97e-1	Accepted	4.67e-1	Accepted
	99%	250	4	7.69e-1	Accepted	4.11e-1	Accepted
EGARCH	95%	250	12	2.13e-2	Accepted	2.50	Accepted
	99%	250	5	1.96e1	Accepted	3.15	Accepted
NA1-GARCH	95%	250	13	2.08e-2	Accepted	1.50e-1	Accepted
	99%	250	7	5.50	**Rejected**	4.05e-1	Accepted

Table A.17: Summary of the Backtesting data for dataset (FTSE20) From 1 May 2013 To 15 April 2014, using four different models when the distribution is assumed to be normal. (T = total number of observations; N = number of exceptions).

Risk Model	Confidence Level	T	N	Kupiec test	Kupiec Outcome	Christ. test	Christ. Outcome
GARCH	95%	250	12	2.13e-2	Accepted	1.22	Accepted
	99%	250	5	1.96e-2	Accepted	2.05e-1	Accepted
TGARCH	95%	250	13	2.08e-2	Accepted	1.43	**Rejected**
	99%	250	2	1.08e-1	Accepted	3.24	**Rejected**
EGARCH	95%	250	1	1.85e1	**Rejected**	1.02	Accepted
	99%	250	0	5.03	**Rejected**	2.05e-1	Accepted
NA1-GARCH	95%	250	11	1.97e-1	Accepted	1.02	Accepted
	99%	250	5	1.96	Accepted	2.05e-1	Accepted

Table A.18: Summary of the Backtesting data for dataset (SP20) From 29 May 2013 To 23 May 2014, using four different models when the distribution is assumed to be normal. (T = total number of observations; N = number of exceptions).

Risk Model	Confidence Level	T	N	Kupiec test	Kupiec Outcome	Christ. test	Christ. Outcome
GARCH	95%	250	12	2.13e-2	Accepted	2.84e-1	Accepted
	99%	250	4	7.69e-1	Accepted	1.31e-1	Accepted
TGARCH	95%	250	28	1.52e1	**Rejected**	4.82	**Rejected**
	99%	250	18	4.11e1	**Rejected**	4.54	**Rejected**
EGARCH	95%	250	15	4.96e-1	Accepted	1.17	Accepted
	99%	250	6	3.56	Accepted	2.42	Accepted
NA1-GARCH	95%	250	12	2.13e-2	Accepted	2.84e-1	Accepted
	99%	250	5	1.96e	Accepted	2.05e-1	Accepted

Table A.19: Summary of the Backtesting data for dataset (FTSE20) From 1 May 2013 To 15 April 2014, using four different models when the distribution is assumed to be student's t-distribution. (T = total number of observations; N = number of exceptions).

Risk Model	Confidence Level	T	N	Kupiec test	Kupiec Outcome	Christ. test	Christ. Outcome
GARCH	95%	250	13	2.08e-2	Accepted	1.43	Accepted
	99%	250	5	1.96	Accepted	2.05	Accepted
TGARCH	95%	250	22	6.26	**Rejected**	6.10e-1	Accepted
	99%	250	14	2.58e1	**Rejected**	4.41	**Rejected**
EGARCH	95%	250	14	1.83e-1	Accepted	1.67	Accepted
	99%	250	5	1.96	Accepted	2.05e-1	Accepted
NA1-GARCH	95%	250	13	2.08e-2	Accepted	1.43	Accepted
	99%	250	6	3.56	Accepted	2.96e-1	Accepted

Table A.20: Summary of the Backtesting data for dataset (SP20) From 29 May 2013 To 23 May 2014, using four different models when the distribution is assumed to be student's t-distribution. (T = total number of observations; N = number of exceptions).

Appendix B

Chapter 4

B.1 Tables of parameter estimates

Datasets	GARCH			NA1-GARCH						NA2-GARCH							EGARCH			
	ω	α	β	ω	α	β	κ	γ	λ	ω	α	β	κ	γ	a	b	ω	α	β	γ
dataset 1	1.47E-04	5.83E-02	1.58E-02	1.70E-05	1.79E-02	6.71E-01	6.21E-04	1.76E-02	7.95E-01	1.50E-05	1.06E-02	7.19E-01	2.41E-02	1.92E-02	1.21E+00	1.55E-03	3.71E-05	-6.70E-02	1.00E+00	3.72E-02
dataset 2	1.34E-04	2.25E-01	3.47E-09	1.83E-05	2.69E-02	9.45E-01	8.08E-03	4.16E-06	1.08E+00	2.13E-05	4.42E-02	3.35E-01	2.31E-06	1.21E-02	1.99E+00	8.74E-03	1.34E-04	-1.69E-02	1.00E+00	1.69E-02
dataset 3	6.59E-06	1.25E-01	8.67E-01	2.37E-06	4.02E-02	8.80E-01	3.73E-02	7.73E-09	9.26E-01	2.97E-06	6.12E-02	5.61E-01	9.84E-03	1.42E-02	1.59E+00	4.05E-01	-4.33E-06	1.69E-01	9.99E-01	-3.37E-02
dataset 4	2.16E-06	7.59E-02	9.20E-01	2.32E-06	3.92E-02	8.77E-01	2.47E-02	1.08E-04	9.27E-01	1.94E-06	6.83E-02	8.28E-01	8.72E-07	1.11E-02	1.11E+00	1.28E-01	-4.41E-04	1.46E-01	1.00E+00	-4.98E-02
dataset 5	3.21E-06	8.58E-16	9.78E-01	1.73E-05	5.21E-02	9.48E-01	1.46E-02	4.03E-09	1.08E+00	1.00E-05	5.51E-02	6.75E-01	7.26E-08	7.70E-03	1.29E+00	2.87E-01	7.88E-05	-4.50E-02	1.00E+00	-1.52E-02
dataset 6	4.05E-05	9.40E-02	6.76E-01	1.36E-05	3.14E-02	2.26E-01	2.20E-02	1.95E-02	3.35E-01	2.22E-05	5.16E-02	3.74E-01	3.66E-08	6.89E-03	1.81E+00	1.39E-01	-1.58E-02	-5.98E-02	9.99E-01	-6.13E-02
dataset 7	7.96E-06	7.07E-02	8.72E-01	6.84E-06	6.07E-02	7.49E-01	1.49E-02	3.19E-03	8.59E-01	6.26E-06	5.56E-02	6.87E-01	4.65E-07	3.35E-02	1.27E+00	2.05E-01	1.29E-04	5.65E-02	1.00E+00	-5.65E-02
dataset 8	2.63E-05	1.80E-02	7.00E-01	2.16E-05	1.67E-02	9.83E-01	3.75E-03	1.33E-02	1.23E+00	1.99E-05	1.37E-02	5.29E-01	1.52E-02	1.94E-03	1.32E+00	2.41E-01	4.58E-06	-3.61E-02	1.00E+00	5.01E-02
dataset 9	1.10E-05	1.22E-01	8.26E-01	1.74E-05	4.57E-02	8.60E-01	3.56E-03	1.38E-02	1.01E+00	6.75E-06	5.27E-02	4.68E-01	1.66E-02	1.45E-08	1.79E+00	2.10E-01	-4.08E-05	1.89E-01	9.98E-01	3.90E-02

Table B.1: Parameter estimates for AstraZeneca datasets

Datasets	GARCH			NA1-GARCH						NA2-GARCH							EGARCH			
	ω	α	β	ω	α	β	κ	γ	λ	ω	α	β	κ	γ	a	b	ω	α	β	γ
dataset 1	1.73E-05	1.08E-01	7.68E-01	1.36E-05	1.31E-01	8.14E-01	4.57E-01	5.40E-01	1.04E+00	1.74E-05	1.17E-01	8.46E-01	3.44E-01	2.85E-01	9.16E-01	1.01E-01	1.03E-05	9.30E-02	3.02E-01	9.50E-04
dataset 2	7.38E-06	1.04E-01	8.65E-01	1.21E-05	1.32E-01	8.68E-01	7.95E-01	4.67E-01	1.05E+00	2.94E-06	5.05E-02	5.02E-01	2.60E-03	9.54E-01	1.77E+00	2.13E-01	-3.84E-06	2.94E-01	9.26E-01	-5.66E-04
dataset 3	6.79E-06	1.17E-01	8.83E-01	3.53E-06	7.61E-02	5.58E-01	8.28E-01	1.59E+00	6.30E-01	5.78E-06	1.21E-01	8.79E-01	4.95E-01	6.27E-01	1.01E+00	1.07E-01	3.56E-04	1.33E-01	1.00E+00	-1.04E-01
dataset 4	3.65E-05	1.72E-01	8.26E-01	1.21E-05	5.79E-02	2.83E-01	1.83E+00	2.44E+00	3.40E-01	1.60E-05	1.55E-01	7.28E-01	5.96E-06	7.13E-07	1.15E+00	7.81E-02	4.45E-04	1.32E-01	1.00E+00	-1.41E-01
dataset 5	1.86E-05	1.48E-01	8.38E-01	1.66E-05	1.36E-01	7.66E-01	8.23E-01	8.01E-01	9.14E-01	1.19E-05	1.17E-01	7.55E-01	7.32E-02	5.46E-01	1.14E+00	7.02E-04	-1.09E-03	8.22E-02	1.00E+00	-1.19E-01
dataset 6	1.69E-05	1.05E-01	8.67E-01	1.60E-05	1.04E-01	8.96E-01	7.10E-01	4.71E-01	1.03E+00	1.46E-05	9.09E-02	7.53E-01	3.50E-01	7.00E-06	1.15E+00	1.01E-01	1.89E-03	1.00E-01	1.00E+00	-8.10E-02
dataset 7	5.46E-06	5.86E-02	9.31E-01	5.28E-07	5.70E-03	8.90E-02	2.19E+00	1.97E+00	9.48E-02	4.21E-06	4.29E-02	6.72E-01	5.81E-04	1.86E-02	1.38E+00	1.34E-01	9.70E-04	5.33E-02	1.00E+00	-7.62E-02
dataset 8	6.43E-06	2.28E-02	9.58E-01	4.93E-06	1.72E-02	6.73E-01	1.99E-01	1.09E+00	7.05E-01	1.04E-05	2.72E-02	9.58E-01	9.32E-01	5.60E-01	9.84E-01	5.63E-02	9.91E-05	-5.90E-02	1.00E+00	-4.18E-03
dataset 9	3.38E-05	8.29E-03	8.55E-01	1.27E-05	8.12E-02	9.19E-01	4.38E-01	7.58E-05	1.03E+00	2.38E-05	4.76E-03	7.08E-01	6.30E-01	9.07E-01	1.23E+00	1.40E-01	-1.18E-05	-8.25E-02	1.00E+00	-1.41E-02

Table B.2: Parameter estimates for Aviva datasets

Datasets	GARCH			NA1-GARCH						NA2-GARCH							EGARCH			
	ω	α	β	ω	α	β	κ	γ	λ	ω	α	β	κ	γ	a	b	ω	α	β	γ
dataset 1	1.03E-05	7.16E-02	8.53E-01	1.40E-05	6.98E-02	8.93E-01	9.81E-02	1.39E-01	1.06E+00	1.01E-05	7.17E-02	9.04E-01	1.01E-01	1.01E-01	9.71E-01	1.01E-01	-2.00E-02	-5.67E-02	9.98E-01	6.01E-02
dataset 2	1.75E-05	7.96E-02	8.02E-01	1.62E-05	7.38E-02	7.44E-01	1.00E-01	1.06E-01	9.28E-01	1.46E-05	6.70E-02	6.78E-01	1.62E-01	9.89E-02	1.19E+00	3.45E-03	1.18E-02	9.69E-02	1.00E+00	-4.02E-02
dataset 3	7.22E-06	1.07E-01	8.78E-01	5.32E-06	7.98E-02	6.54E-01	1.14E-01	1.94E-01	7.45E-01	7.22E-06	1.07E-01	8.77E-01	9.61E-02	7.19E-09	1.00E+00	4.44E-02	-1.97E-01	1.90E-01	9.75E-01	-6.54E-02
dataset 4	1.26E-05	8.60E-02	8.83E-01	8.25E-06	5.38E-02	4.89E-01	1.36E-06	1.06E-01	5.61E-01	9.55E-06	5.61E-02	4.24E-01	4.97E-03	1.46E-01	2.00E+00	3.07E-07	-8.79E-02	1.75E-01	9.89E-01	-5.22E-02
dataset 5	1.26E-05	1.18E-01	8.48E-01	1.08E-05	9.63E-02	7.18E-01	1.47E-03	1.59E-01	8.44E-01	1.26E-05	1.18E-01	8.47E-01	3.66E-08	2.12E-02	1.00E+00	2.35E-07	-2.29E-01	7.94E-02	9.72E-01	-1.30E-01
dataset 6	1.17E-05	1.13E-01	8.59E-01	1.02E-07	8.58E-04	6.18E-03	3.24E-01	2.69E-01	7.06E-03	9.52E-06	9.50E-02	8.11E-01	1.33E-05	1.53E-01	1.08E+00	1.50E-07	2.09E-02	2.15E-01	1.00E+00	-9.32E-02
dataset 7	6.62E-06	6.01E-02	9.10E-01	7.07E-06	6.40E-02	9.36E-01	1.28E-01	1.12E-01	1.03E+00	6.36E-06	5.81E-02	8.91E-01	3.43E-07	1.58E-01	1.02E+00	9.89E-02	4.08E-02	2.13E-02	1.01E+00	-1.01E-01
dataset 8	5.74E-05	1.90E-01	3.77E-01	3.64E-05	1.25E-01	4.30E-01	3.67E-02	8.19E-04	8.32E-01	1.13E-05	4.55E-02	3.67E-01	1.91E-01	1.08E-01	2.00E+00	1.96E-03	4.39E-04	-5.79E-02	1.00E+00	3.53E-03
dataset 9	3.31E-07	1.34E-02	9.87E-01	7.92E-12	1.10E-02	9.83E-01	4.44E-01	1.86E-01	9.90E-01	8.95E-06	4.83E-02	9.52E-01	1.22E-03	1.15E-01	9.32E-01	2.70E-01	-1.90E-01	1.86E-02	9.79E-01	-6.31E-02

Table B.3: Parameter estimates for BP datasets

Datasets	GARCH			NA1-GARCH						NA2-GARCH							EGARCH			
	ω	α	β	ω	α	β	κ	γ	λ	ω	α	β	κ	γ	a	b	ω	α	β	γ
dataset 1	2.06E-05	8.01E-03	8.08E-01	1.47E-05	3.91E-02	9.61E-01	7.80E-04	7.02E-04	1.13E+00	1.73E-05	1.57E-02	8.44E-01	1.17E-11	9.22E-04	9.88E-01	2.97E-01	9.96E-06	1.24E-02	1.00E+00	3.28E-03
dataset 2	9.17E-05	1.47E-01	8.77E-02	2.86E-05	4.60E-02	2.74E-02	9.89E-04	5.61E-04	3.12E-01	1.90E-05	3.08E-02	9.68E-01	4.26E-04	1.75E-03	8.61E-01	2.21E-04	2.57E-03	-3.63E-02	1.00E+00	-4.70E-03
dataset 3	1.56E-05	1.34E-01	8.14E-01	5.75E-06	6.43E-02	5.56E-01	2.74E-03	1.92E-08	6.40E-01	7.63E-06	4.27E-02	5.16E-01	7.40E-08	2.36E-03	1.69E+00	2.81E-01	-2.87E-04	1.64E-01	9.99E-01	-6.72E-02
dataset 4	6.37E-06	7.96E-02	8.99E-01	3.89E-06	5.67E-02	7.29E-01	2.37E-03	7.02E-13	7.99E-01	3.08E-06	4.09E-02	4.95E-01	1.23E-03	2.01E-03	1.83E+00	1.71E-01	3.53E-05	-4.40E-02	1.00E+00	-4.74E-02
dataset 5	3.67E-05	1.17E-01	6.81E-01	1.11E-05	3.52E-02	2.05E-01	1.90E-03	1.39E-03	3.01E-01	1.83E-05	5.54E-02	3.48E-01	4.20E-04	2.92E-03	1.98E+00	1.14E-05	-1.09E-04	-5.96E-02	1.00E+00	-7.92E-02
dataset 6	2.15E-05	1.19E-01	7.33E-01	3.91E-06	2.16E-02	1.33E-01	1.77E-03	1.64E-03	1.82E-01	9.97E-06	4.84E-02	4.50E-01	2.00E-03	1.46E-05	1.76E+00	2.41E-01	1.50E-04	-5.45E-02	1.00E+00	-1.31E-02
dataset 7	2.36E-05	1.36E-01	6.61E-01	1.08E-05	4.61E-02	3.07E-01	1.45E-06	6.15E-07	4.50E-01	1.65E-05	1.04E-01	8.96E-01	7.06E-04	5.48E-05	8.76E-01	5.98E-06	6.74E-05	5.00E-02	1.00E+00	-1.87E-02
dataset 8	5.81E-06	2.45E-02	9.07E-01	1.10E-05	4.54E-02	9.12E-01	2.45E-10	3.06E-03	1.09E+00	4.29E-06	1.83E-02	6.23E-01	7.20E-04	3.45E-05	1.45E+00	5.37E-01	4.80E-04	-1.24E-02	1.00E+00	3.44E-02
dataset 9	1.62E-05	3.97E-02	8.04E-01	1.86E-05	4.92E-02	9.51E-01	1.15E-03	3.19E-05	1.18E+00	1.73E-05	4.32E-02	9.56E-01	1.49E-03	7.85E-04	8.56E-01	9.48E-07	3.19E-05	-1.39E-02	1.00E+00	-2.36E-02

Table B.4: Parameter estimates for GlaxoSmithKline datasets

Datasets	GARCH			NA1-GARCH						NA2-GARCH							EGARCH			
	ω	α	β	ω	α	β	κ	γ	λ	ω	α	β	κ	γ	a	b	ω	α	β	γ
dataset 1	7.96E-06	2.28E-02	9.06E-01	1.16E-06	3.55E-02	9.64E-01	1.21E-03	1.31E-03	1.01E+00	1.57E-06	2.54E-02	9.36E-01	9.00E-05	1.18E-03	1.03E+00	1.71E-01	1.01E-07	3.33E-02	1.00E+00	4.48E-02
dataset 2	4.85E-06	8.80E-02	8.87E-01	2.86E-06	2.21E-02	3.28E-01	9.73E-10	2.26E-03	3.68E-01	1.17E-06	5.20E-02	5.26E-01	1.93E-03	7.30E-06	1.72E+00	2.47E-01	-5.43E-06	1.19E-01	9.99E-01	-3.02E-02
dataset 3	3.73E-06	1.10E-01	8.90E-01	1.23E-06	5.13E-02	3.67E-01	1.38E-03	1.74E-03	4.12E-01	9.43E-07	5.27E-02	4.96E-01	4.27E-04	6.13E-04	1.84E+00	1.45E-01	2.28E-06	-4.57E-02	1.00E+00	-1.16E-01
dataset 4	4.32E-20	9.26E-02	9.07E-01	1.11E-05	1.02E-01	5.76E-01	5.06E-04	5.61E-10	6.69E-01	4.68E-09	5.88E-02	6.40E-01	4.34E-03	7.58E-06	1.44E+00	1.75E-06	1.45E-06	-5.67E-02	1.00E+00	-1.28E-01
dataset 5	2.60E-05	4.53E-02	9.23E-01	1.68E-06	6.10E-02	3.79E-01	1.56E-08	2.09E-05	4.35E-01	5.73E-08	4.20E-02	7.17E-01	2.38E-06	6.05E-03	1.31E+00	6.12E-02	-5.05E-06	1.55E-01	1.00E+00	-3.42E-02
dataset 6	4.70E-05	1.07E-01	8.48E-01	2.97E-06	6.83E-03	5.40E-02	4.12E-05	5.40E-05	6.36E-02	3.00E-06	3.52E-02	4.99E-01	3.78E-07	2.12E-03	1.87E+00	1.27E-01	-6.42E-03	5.55E-02	9.99E-01	-6.88E-02
dataset 7	1.37E-05	4.99E-02	9.35E-01	2.74E-06	2.52E-02	5.30E-01	1.46E-03	1.33E-06	5.57E-01	9.67E-07	3.18E-02	6.21E-01	5.23E-04	1.77E-03	1.53E+00	1.67E-01	3.25E-06	8.98E-02	1.00E+00	-4.05E-02
dataset 8	1.24E-06	2.85E-15	9.93E-01	1.16E-06	4.23E-02	9.58E-01	1.11E-03	1.04E-03	1.00E+00	1.01E-06	5.35E-04	9.06E-01	1.84E-03	4.07E-04	1.10E+00	2.28E-01	1.85E-06	-5.48E-02	1.00E+00	-1.74E-02
dataset 9	1.04E-06	8.10E-15	9.93E-01	1.80E-06	3.31E-02	9.62E-01	1.52E-03	1.48E-08	1.00E+00	1.28E-06	2.52E-02	9.75E-01	1.53E-03	3.44E-05	9.94E-01	1.94E-01	2.28E-06	-6.15E-02	1.00E+00	-2.04E-03

Table B.5: Parameter estimates for Lloyds Bank datasets

Datasets	GARCH			NA1-GARCH						NA2-GARCH							EGARCH			
	ω	α	β	ω	α	β	κ	γ	λ	ω	α	β	κ	γ	a	b	ω	α	β	γ
dataset 1	1.01E-05	6.20E-02	8.91E-01	1.16E-05	6.27E-02	9.33E-01	1.07E+00	1.29E+00	1.06E+00	9.48E-06	6.10E-02	9.38E-01	9.90E-01	8.63E-01	9.61E-01	8.22E-02	1.02E-05	6.17E-02	2.55E-01	9.59E-03
dataset 2	1.54E-05	7.43E-02	8.72E-01	1.27E-05	1.01E-01	8.99E-01	1.28E+00	4.54E-01	1.04E+00	9.54E-06	7.98E-02	9.20E-01	4.39E-01	4.14E-01	9.70E-01	6.64E-02	-6.72E-06	9.03E-02	9.99E-01	-2.35E-02
dataset 3	1.91E-06	1.28E-02	9.87E-01	9.88E-06	8.53E-02	9.15E-01	9.29E-01	8.26E-01	1.01E+00	8.25E-06	7.51E-02	9.25E-01	5.25E-01	1.02E+00	9.90E-01	1.18E-01	-1.07E-05	2.34E-01	9.51E-01	-9.77E-03
dataset 4	1.53E-06	6.46E-02	9.35E-01	1.55E-06	6.44E-02	9.35E-01	1.41E+00	7.83E-01	9.99E-01	1.29E-06	5.41E-02	7.87E-01	3.01E+00	7.08E-01	1.19E+00	2.88E-02	-5.52E-05	1.04E-01	1.00E+00	-6.72E-02
dataset 5	2.94E-06	3.43E-02	9.48E-01	2.28E-06	3.39E-02	9.09E-01	6.69E-01	4.18E-03	9.56E-01	2.50E-06	3.81E-02	8.01E-01	2.35E+00	1.35E+00	1.17E+00	2.60E-07	-2.69E-05	1.41E-01	1.00E+00	5.66E-03
dataset 6	1.70E-05	6.49E-02	8.36E-01	5.62E-06	2.06E-02	2.09E-01	7.80E+00	4.23E-01	2.54E-01	1.83E-05	7.04E-02	9.27E-01	1.15E+00	1.32E+00	9.03E-01	7.85E-02	7.07E-05	1.01E-01	9.99E-01	-7.47E-02
dataset 7	1.28E-05	5.09E-02	8.57E-01	1.03E-05	4.46E-02	9.50E-01	2.50E+00	6.19E-01	1.07E+00	1.20E-05	4.78E-02	7.94E-01	1.96E+00	4.73E-02	1.08E+00	6.04E-02	3.14E-05	-4.86E-02	1.00E+00	3.45E-02
dataset 8	4.68E-05	3.98E-02	6.42E-01	2.41E-05	2.05E-02	3.29E-01	1.56E+00	7.71E-01	5.12E-01	3.43E-05	3.05E-02	5.66E-01	3.00E+00	1.07E-01	1.19E+00	3.66E-01	-6.94E-05	8.02E-01	9.80E-01	-5.58E-02
dataset 9	3.79E-05	3.23E-02	7.83E-01	2.16E-05	2.32E-02	5.00E-01	2.79E+00	1.98E-06	6.27E-01	2.10E-05	2.17E-02	6.52E-01	1.30E+00	2.00E+00	1.29E+00	6.37E-03	-2.68E-05	-5.93E-02	1.00E+00	3.00E-02

Table B.6: Parameter estimates for Vodafone datasets

Datasets	GARCH			NA1-GARCH						NA2-GARCH							EGARCH			
	ω	α	β	ω	α	β	κ	γ	λ	ω	α	β	κ	γ	a	b	ω	α	β	γ
dataset 1	1.07E-05	7.69E-02	8.54E-01	8.27E-06	6.00E-02	6.65E-01	1.76E-08	4.28E-01	7.78E-01	8.42E-06	6.28E-02	9.33E-01	9.38E-02	1.04E-01	9.52E-01	9.19E-09	-5.64E-01	2.06E-01	9.35E-01	-9.62E-02
dataset 2	1.75E-05	8.53E-02	8.25E-01	1.49E-05	7.43E-02	7.37E-01	3.25E-02	1.33E-10	8.88E-01	1.22E-05	6.00E-02	5.80E-01	4.05E-02	5.40E-02	1.42E+00	2.54E-01	-4.69E-01	6.44E-02	9.46E-01	-9.25E-02
dataset 3	1.43E-05	1.60E-01	8.31E-01	4.50E-06	5.04E-02	2.62E-01	8.21E-02	6.70E-01	3.15E-01	6.28E-06	7.21E-02	5.15E-01	1.02E-01	1.17E-03	1.69E+00	1.45E-02	-1.92E-01	1.60E-01	9.75E-01	-1.34E-01
dataset 4	2.63E-05	1.40E-01	8.41E-01	1.95E-05	1.03E-01	6.24E-01	2.23E-01	1.34E-02	7.41E-01	9.61E-06	6.01E-02	4.77E-01	2.16E-01	2.32E-03	1.83E+00	1.76E-03	-1.12E-01	5.67E-02	9.85E-01	-1.47E-01
dataset 5	4.47E-06	6.95E-02	9.15E-01	3.30E-06	5.11E-02	6.73E-01	2.58E-01	1.88E-01	7.35E-01	3.40E-06	5.29E-02	6.96E-01	1.71E-01	6.01E-02	1.31E+00	1.29E-01	-9.59E-02	1.20E-01	9.88E-01	-8.44E-02
dataset 6	6.40E-06	1.05E-01	8.85E-01	2.91E-06	4.83E-02	4.13E-01	2.40E-01	2.15E-01	4.66E-01	7.26E-06	7.74E-02	4.09E-01	2.99E-02	8.51E-02	2.00E+00	1.23E-05	-1.35E-01	2.17E-01	9.82E-01	-8.69E-02
dataset 7	9.55E-06	7.62E-02	9.02E-01	8.51E-06	6.79E-02	8.02E-01	3.80E-02	1.25E-04	8.89E-01	7.46E-06	5.95E-02	7.05E-01	2.58E-02	1.40E-01	1.28E+00	5.45E-03	-3.13E-01	1.71E-01	9.59E-01	-1.24E-01
dataset 8	4.42E-06	5.20E-02	9.23E-01	3.06E-06	3.60E-02	6.38E-01	3.61E-07	3.00E-01	6.91E-01	4.20E-06	4.96E-02	8.86E-01	6.96E-02	9.80E-02	1.04E+00	2.57E-01	-2.81E-01	9.55E-02	9.66E-01	-1.08E-01
dataset 9	6.05E-06	3.74E-02	9.22E-01	6.37E-06	3.84E-02	8.79E-01	1.39E-03	8.77E-06	9.61E-01	6.77E-06	3.64E-02	7.92E-01	6.99E-07	8.19E-02	1.14E+00	3.69E-01	-3.58E-01	9.47E-02	9.59E-01	-5.08E-03

Table B.7: Parameter estimates for Allianz datasets

Datasets	GARCH			NA1-GARCH						NA2-GARCH							EGARCH			
	ω	α	β	ω	α	β	κ	γ	λ	ω	α	β	κ	γ	a	b	ω	α	β	γ
dataset 1	1.00E-05	2.09E-02	9.16E-01	1.20E-05	2.67E-02	9.67E-01	1.20E-02	1.09E-02	1.07E+00	1.01E-05	1.90E-02	9.45E-01	1.01E-02	9.87E-03	9.73E-01	9.80E-02	1.63E-05	1.42E-02	1.00E+00	-2.09E-02
dataset 2	2.36E-05	6.19E-02	8.37E-01	1.35E-05	5.63E-02	8.65E-01	1.88E-02	8.83E-03	9.76E-01	1.41E-05	3.42E-02	4.67E-01	2.76E-02	2.94E-03	1.78E+00	8.78E-02	6.22E-05	7.40E-02	9.99E-01	-5.13E-02
dataset 3	4.27E-06	3.71E-02	9.63E-01	7.45E-06	6.57E-02	9.15E-01	1.13E-02	6.63E-03	9.80E-01	5.87E-06	5.08E-02	6.94E-01	7.42E-03	9.28E-03	1.34E+00	9.15E-02	-2.37E-04	-5.78E-02	9.99E-01	-7.27E-02
dataset 4	1.93E-05	1.14E-01	8.70E-01	6.73E-06	4.36E-02	8.56E-01	2.68E-02	9.19E-09	9.10E-01	5.91E-06	3.73E-02	4.59E-01	4.05E-03	1.58E-02	1.99E+00	1.30E-02	3.25E-04	-6.09E-02	1.00E+00	-9.10E-02
dataset 5	8.43E-06	1.14E-10	9.65E-01	9.08E-06	2.41E-02	9.76E-01	1.18E-02	1.39E-02	1.04E+00	9.82E-06	1.66E-02	9.83E-01	9.16E-03	5.84E-03	9.63E-01	9.90E-02	5.68E-05	-4.91E-02	1.00E+00	-1.65E-01
dataset 6	1.88E-05	9.53E-02	8.21E-01	1.57E-05	9.21E-02	8.73E-01	1.45E-01	4.86E-03	1.04E+00	1.17E-05	3.54E-02	4.46E-01	7.43E-03	3.37E-04	1.86E+00	1.17E-01	-1.17E-04	8.18E-02	1.00E+00	-1.07E-01
dataset 7	1.10E-05	5.81E-02	8.81E-01	8.17E-06	4.33E-02	6.57E-01	1.58E-02	7.85E-02	7.46E-01	7.53E-06	4.00E-02	6.04E-01	3.46E-04	3.52E-03	1.46E+00	1.73E-01	4.81E-05	-4.17E-02	1.00E+00	-3.04E-02
dataset 8	1.55E-04	6.66E-02	1.03E-01	4.37E-05	1.80E-02	3.01E-02	4.78E-02	4.75E-03	2.76E-01	2.20E-05	1.54E-02	6.02E-01	3.91E-03	6.82E-03	1.36E+00	5.26E-02	2.93E-05	-2.39E-02	1.00E+00	-4.28E-02
dataset 9	7.84E-06	2.37E-02	9.31E-01	1.09E-05	2.88E-02	9.71E-01	1.17E-02	7.73E-03	1.06E+00	7.74E-06	2.46E-02	9.75E-01	1.35E-02	7.57E-03	9.57E-01	8.85E-02	5.82E-05	-5.00E-02	1.00E+00	-1.08E-01

Table B.8: Parameter estimates for Anheuser-Busch datasets

Datasets	GARCH			NA1-GARCH						NA2-GARCH							EGARCH			
	ω	α	β	ω	α	β	κ	γ	λ	ω	α	β	κ	γ	a	b	ω	α	β	γ
dataset 1	1.02E-05	3.14E-02	8.69E-01	1.15E-05	3.26E-02	9.49E-01	1.00E-01	1.06E-01	1.10E+00	1.03E-05	3.04E-02	9.37E-01	9.75E-02	9.76E-02	9.36E-01	1.97E-01	1.77E-05	-4.15E-02	1.00E+00	-6.07E-02
dataset 2	1.01E-05	1.06E-01	8.30E-01	5.20E-06	5.61E-02	4.39E-01	2.46E-03	1.80E-01	5.28E-01	8.24E-06	8.74E-02	6.88E-01	1.86E-01	9.68E-06	1.21E+00	4.00E-01	3.22E-07	1.86E-01	9.99E-01	-4.67E-02
dataset 3	9.66E-06	1.66E-01	8.29E-01	3.24E-06	5.56E-02	2.78E-01	2.27E-01	4.08E-02	3.35E-01	4.71E-06	8.10E-02	4.86E-01	3.41E-06	5.48E-02	1.75E+00	2.96E-05	-5.41E-05	1.61E-01	1.00E+00	-1.25E-01
dataset 4	1.45E-05	1.45E-01	8.51E-01	2.14E-06	2.09E-02	1.24E-01	2.21E-01	3.00E-01	1.45E-01	6.02E-06	6.37E-02	4.69E-01	1.52E-01	1.03E-01	1.87E+00	1.34E-01	-1.98E-05	2.03E-01	1.00E+00	-1.52E-01
dataset 5	2.03E-05	1.68E-01	8.24E-01	7.17E-06	5.92E-02	2.91E-01	5.59E-02	9.77E-02	3.53E-01	1.02E-05	6.47E-02	4.58E-01	3.61E-02	3.77E-04	1.88E+00	1.68E-03	-1.29E-04	2.09E-01	1.00E+00	-1.65E-01
dataset 6	3.39E-05	1.46E-01	8.18E-01	1.90E-05	8.15E-02	4.58E-01	5.73E-02	6.04E-02	5.60E-01	1.21E-05	6.90E-02	4.98E-01	1.03E-01	8.55E-02	1.72E+00	3.54E-03	1.05E-04	-7.66E-02	1.00E+00	-1.45E-01
dataset 7	2.02E-05	5.30E-02	9.13E-01	1.12E-07	3.76E-04	6.59E-03	2.48E-01	9.34E-02	7.08E-03	1.43E-05	3.77E-02	6.55E-01	3.22E-07	1.36E-01	1.40E+00	1.28E-01	7.76E-05	-6.61E-02	1.00E+00	-9.08E-02
dataset 8	9.56E-06	6.64E-02	9.11E-01	4.85E-08	3.25E-04	5.09E-03	4.40E-01	1.23E-01	5.43E-03	7.26E-06	5.07E-02	6.92E-01	4.62E-02	1.14E-01	1.32E+00	2.02E-01	7.34E-07	-3.65E-02	1.00E+00	-3.84E-02
dataset 9	2.15E-05	7.10E-02	8.38E-01	1.48E-07	5.25E-04	6.63E-03	2.73E-01	1.56E-01	7.68E-03	1.32E-05	4.37E-02	5.18E-01	6.63E-02	4.64E-02	1.62E+00	2.15E-01	-2.41E-04	3.97E-02	1.00E+00	-4.80E-02

Table B.9: Parameter estimates for Banco Santander datasets

Datasets	GARCH			NA1-GARCH						NA2-GARCH							EGARCH			
	ω	α	β	ω	α	β	κ	γ	λ	ω	α	β	κ	γ	a	b	ω	α	β	γ
dataset 1	6.75E-06	6.54E-02	8.96E-01	5.11E-06	3.15E-02	9.68E-01	1.89E-02	1.64E-02	1.03E+00	5.11E-06	5.08E-02	7.90E-01	8.93E-03	9.17E-03	1.15E+00	2.75E-06	8.91E-03	5.94E-02	2.77E-01	1.53E-02
dataset 2	9.16E-05	2.74E-01	3.28E-01	1.92E-05	5.76E-02	6.90E-02	9.23E-04	3.12E-02	2.10E-01	8.52E-06	5.31E-02	6.36E-01	1.86E-03	1.47E-02	1.38E+00	1.07E-01	-2.33E-01	1.01E-01	9.72E-01	-9.81E-02
dataset 3	1.09E-05	1.02E-01	8.79E-01	3.62E-06	3.23E-02	2.93E-01	1.10E-02	2.56E-02	3.32E-01	5.95E-06	5.65E-02	4.90E-01	5.16E-03	8.47E-07	1.80E+00	3.81E-02	-7.99E-02	2.34E-01	9.88E-01	-2.93E-02
dataset 4	2.55E-05	1.19E-01	8.39E-01	1.32E-05	5.53E-02	4.17E-01	1.13E-03	3.50E-02	4.95E-01	1.41E-05	5.14E-02	5.86E-01	5.00E-03	2.67E-02	1.50E+00	2.22E-05	-3.15E-01	1.46E-01	9.58E-01	-1.26E-01
dataset 5	2.62E-05	8.89E-02	8.28E-01	2.13E-05	4.49E-02	3.79E-01	5.03E-05	2.85E-02	4.91E-01	1.42E-05	3.87E-02	7.19E-01	4.79E-03	1.64E-02	1.24E+00	2.44E-06	-3.08E-01	9.04E-02	9.62E-01	-6.76E-02
dataset 6	9.96E-07	3.36E-02	9.66E-01	5.26E-06	4.76E-02	9.52E-01	1.15E-02	2.91E-02	1.02E+00	1.16E-05	7.93E-02	9.21E-01	1.29E-02	2.10E-02	9.67E-01	3.23E-02	-8.31E-02	5.57E-02	9.90E-01	-9.68E-03
dataset 7	8.39E-06	4.29E-02	9.34E-01	4.31E-06	2.30E-02	4.91E-01	3.15E-03	3.31E-02	5.26E-01	8.58E-06	4.41E-02	9.56E-01	8.34E-03	5.69E-03	9.77E-01	1.33E-01	-9.58E-02	-5.33E-02	9.88E-01	-1.09E-01
dataset 8	3.70E-18	6.83E-16	9.98E-01	7.36E-06	3.41E-02	9.66E-01	1.47E-02	1.24E-02	1.04E+00	4.83E-06	2.42E-02	9.74E-01	1.19E-02	1.33E-02	9.78E-01	1.85E-05	1.29E-01	-3.58E-02	1.02E+00	-1.34E-02
dataset 9	7.16E-06	1.86E-02	9.45E-01	1.54E-05	2.57E-02	9.74E-01	8.44E-03	1.27E-02	1.08E+00	7.50E-06	1.94E-02	9.81E-01	4.60E-03	1.41E-02	9.63E-01	8.07E-02	-4.27E-01	2.30E-02	9.49E-01	-1.01E-01

Table B.10: Parameter estimates for Bayer datasets

Datasets	GARCH			NA1-GARCH						NA2-GARCH							EGARCH			
	ω	α	β	ω	α	β	κ	γ	λ	ω	α	β	κ	γ	a	b	ω	α	β	γ
dataset 1	1.03E-05	6.17E-02	8.67E-01	1.42E-06	4.40E-02	9.23E-01	1.68E-03	1.12E-03	9.76E-01	9.05E-07	3.34E-02	9.67E-01	1.07E-03	7.95E-04	9.94E-01	2.82E-01	-6.04E-01	1.13E-01	9.31E-01	-5.55E-02
dataset 2	1.58E-05	1.14E-01	8.09E-01	2.44E-06	6.03E-02	9.37E-01	9.35E-07	1.64E-03	1.00E+00	3.62E-06	3.70E-02	4.64E-01	5.46E-04	5.49E-04	1.93E+00	4.10E-04	-1.18E-01	5.89E-02	9.86E-01	-8.99E-02
dataset 3	1.33E-05	1.89E-01	8.11E-01	9.69E-10	1.08E-05	4.37E-05	1.15E-03	3.15E-03	5.36E-05	1.79E-06	4.88E-02	4.55E-01	7.44E-04	6.08E-04	2.00E+00	7.18E-07	-4.72E-02	1.43E-01	9.94E-01	-1.38E-01
dataset 4	5.25E-11	6.25E-02	9.38E-01	4.29E-06	7.45E-02	4.40E-01	7.93E-04	2.15E-07	5.06E-01	1.70E-06	4.61E-02	4.61E-01	1.18E-03	1.38E-04	1.99E+00	2.63E-04	4.50E-03	9.50E-02	1.00E+00	-9.37E-02
dataset 5	2.77E-19	4.43E-02	9.53E-01	6.32E-07	6.19E-07	9.42E-02	5.14E-04	2.44E-03	9.57E-02	1.79E-06	2.63E-02	5.73E-01	1.16E-03	2.33E-05	1.66E+00	3.38E-01	3.48E-03	8.30E-02	1.00E+00	-3.18E-02
dataset 6	9.80E-07	7.25E-02	9.28E-01	1.97E-06	3.08E-02	2.60E-01	2.44E-11	2.63E-03	2.92E-01	1.17E-06	4.16E-02	4.87E-01	1.43E-03	4.93E-05	1.90E+00	3.05E-02	8.61E-03	1.58E-01	1.00E+00	-3.02E-02
dataset 7	9.32E-06	6.17E-02	9.28E-01	5.80E-06	3.84E-02	5.78E-01	1.46E-03	1.94E-03	6.22E-01	1.31E-06	3.91E-02	6.13E-01	1.22E-05	1.24E-03	1.54E+00	8.57E-02	-1.78E-02	9.02E-02	9.97E-01	-5.16E-02
dataset 8	1.31E-16	3.61E-02	9.61E-01	7.20E-07	4.49E-02	9.55E-01	1.48E-03	1.06E-03	1.00E+00	4.44E-07	3.95E-02	9.61E-01	1.62E-03	3.07E-03	9.96E-01	2.56E-01	8.03E-03	-5.87E-02	1.00E+00	-7.45E-02
dataset 9	7.51E-06	5.04E-02	9.19E-01	4.78E-06	1.45E-02	1.66E-01	1.82E-03	7.43E-06	2.00E-01	3.05E-06	4.52E-02	4.66E-01	7.54E-04	4.78E-04	1.92E+00	2.44E-04	7.45E-02	1.64E-01	1.01E+00	1.96E-02

Table B.11: Parameter estimates for Deutsche Bank datasets

Datasets	GARCH			NA1-GARCH						NA2-GARCH							EGARCH			
	ω	α	β	ω	α	β	κ	γ	λ	ω	α	β	κ	γ	a	b	ω	α	β	γ
dataset 1	1.03E-05	3.05E-02	8.83E-01	1.29E-05	3.99E-02	8.42E-01	9.88E-03	1.32E-02	9.90E-01	1.03E-05	3.09E-02	8.88E-01	1.02E-02	1.07E-02	9.96E-01	3.11E-01	1.94E-05	2.15E-07	1.00E+00	1.69E-02
dataset 2	8.98E-06	6.90E-02	8.71E-01	6.27E-06	4.82E-02	6.08E-01	1.51E-02	2.06E-02	6.98E-01	8.45E-06	6.48E-02	8.19E-01	1.38E-02	4.69E-03	1.06E+00	1.79E-01	-6.84E-05	1.22E-06	1.00E+00	-8.23E-03
dataset 3	7.07E-06	1.36E-01	8.57E-01	7.50E-06	3.90E-02	9.49E-01	1.21E-06	2.67E-02	1.01E+00	5.43E-06	6.14E-02	5.17E-01	1.60E-02	5.64E-06	1.69E+00	3.09E-01	9.57E-02	-9.54E-04	1.01E+00	-2.09E-02
dataset 4	1.18E-05	1.15E-01	8.66E-01	1.50E-06	1.72E-02	1.55E-01	1.19E-03	8.93E-03	1.75E-01	6.40E-06	6.07E-02	4.98E-01	2.76E-02	6.66E-08	1.76E+00	2.45E-01	-8.28E-02	3.64E-04	9.91E-01	-1.55E-01
dataset 5	3.82E-06	2.44E-02	9.57E-01	4.21E-06	4.00E-02	6.18E-01	3.58E-03	2.89E-02	6.74E-01	2.59E-06	2.11E-02	6.88E-01	6.77E-03	1.48E-02	1.39E+00	3.09E-01	6.54E-06	2.11E-07	1.00E+00	-7.83E-02
dataset 6	1.42E-05	1.08E-01	8.30E-01	8.01E-06	6.09E-02	4.69E-01	1.06E-02	3.38E-02	5.65E-01	7.87E-06	6.42E-02	5.51E-01	2.54E-02	6.74E-03	1.54E+00	4.25E-01	1.48E-04	1.69E-06	1.00E+00	-8.17E-02
dataset 7	1.29E-05	5.41E-02	8.84E-01	8.28E-06	3.35E-02	5.15E-01	3.45E-07	2.02E-02	5.88E-01	7.38E-06	3.04E-02	4.81E-01	7.86E-03	1.98E-02	1.83E+00	1.72E-01	-4.85E-04	-3.52E-06	1.00E+00	-7.28E-02
dataset 8	2.12E-05	8.56E-02	7.66E-01	1.32E-05	5.32E-02	4.76E-01	1.29E-02	9.14E-03	6.21E-01	1.57E-05	6.37E-02	5.69E-01	7.87E-03	7.11E-03	1.34E+00	1.63E-01	2.80E-05	1.25E-07	1.00E+00	-4.83E-02
dataset 9	9.96E-06	1.12E-01	8.24E-01	4.90E-06	5.53E-02	4.06E-01	1.82E-02	1.80E-02	4.92E-01	7.80E-06	8.89E-02	6.61E-01	8.62E-04	1.26E-02	1.25E+00	8.79E-05	4.06E-04	-1.11E-05	1.00E+00	-4.33E-02

Table B.12: Parameter estimates for Total datasets

B.2 Tables of out-of-sample errors and in-sample model comparison

In each table and for each dataset, the lower of three errors or AIC values is highlighted in boldface.

Datasets	MAE				RMSE				AIC			
	GARCH	NA1-GARCH	NA2-GARCH	EGARCH	GARCH	NA1-GARCH	NA2-GARCH	EGARCH	GARCH	NA1-GARCH	NA2-GARCH	EGARCH
dataset 1	2.80E-03	**2.58E-03**	2.71E-03	1.19E-02	3.59E-03	**3.22E-03**	3.36E-03	1.71E-02	**-3.90E+03**	-3.86E+03	-3.86E+03	-3.89E+03
dataset 2	8.41E-03	7.56E-03	**6.84E-03**	1.19E-02	1.18E-02	1.02E-02	**9.75E-03**	1.71E-02	**-3.86E+03**	-3.84E+03	-3.84E+03	-3.84E+03
dataset 3	2.30E-03	**1.82E-03**	1.99E-03	2.10E-03	2.97E-03	**2.16E-03**	2.55E-03	3.09E-03	-3.57E+03	-3.54E+03	-3.56E+03	**-3.57E+03**
dataset 4	2.31E-03	4.88E-03	2.31E-03	**2.22E-03**	**3.45E-03**	8.41E-03	**3.45E-03**	3.56E-03	-3.46E+03	-3.44E+03	-3.45E+03	**-3.47E+03**
dataset 5	6.33E-03	2.73E-03	**2.40E-03**	1.53E-01	7.72E-03	3.34E-03	**3.00E-03**	2.35E-01	**-3.78E+03**	-3.75E+03	-3.75E+03	-3.78E+03
dataset 6	3.36E-03	3.36E-03	**3.35E-03**	7.29E-02	3.70E-03	3.70E-03	**3.69E-03**	1.13E-01	-3.85E+03	-3.84E+03	-3.84E+03	**-3.88E+03**
dataset 7	1.66E-03	1.66E-03	**1.65E-03**	1.72E-03	1.94E-03	1.94E-03	**1.94E-03**	2.02E-03	**-3.98E+03**	-3.97E+03	-3.97E+03	-3.98E+03
dataset 8	4.74E-03	**4.69E-03**	4.74E-03	6.29E-03	8.78E-03	**8.63E-03**	8.78E-03	8.89E-03	-4.14E+03	-4.13E+03	-4.13E+03	**-4.15E+03**
dataset 9	1.79E-03	1.86E-03	**1.67E-03**	1.99E-03	2.48E-03	2.55E-03	**2.36E-03**	2.48E-03	**-3.90E+03**	-3.87E+03	-3.89E+03	-3.88E+03

Table B.13: MAE, RMSE and AIC for GARCH, NA1-GARCH, NA2-GARCH and EGARCH models for AstraZeneca datasets.

Datasets	MAE				RMSE				AIC			
	GARCH	NA1-GARCH	NA2-GARCH	EGARCH	GARCH	NA1-GARCH	NA2-GARCH	EGARCH	GARCH	NA1-GARCH	NA2-GARCH	EGARCH
dataset 1	2.99E-03	**2.66E-03**	2.94E-03	2.02E-01	3.98E-03	**3.53E-03**	3.92E-03	2.07E-01	**-4.01E+03**	-3.99E+03	-3.98E+03	-9.60E+01
dataset 2	5.55E-03	6.86E-03	**4.83E-03**	2.02E-01	8.03E-03	9.84E-03	**7.03E-03**	2.07E-01	**-3.77E+03**	-3.76E+03	-3.76E+03	-1.54E+03
dataset 3	**6.06E-03**	6.19E-03	6.23E-03	7.69E-03	1.02E-02	**1.02E-02**	1.04E-02	1.12E-02	-3.19E+03	-3.19E+03	-3.18E+03	**-3.20E+03**
dataset 4	4.91E-03	4.89E-03	3.98E-03	**3.04E-03**	5.85E-03	5.82E-03	5.09E-03	**3.85E-03**	-2.73E+03	-2.72E+03	-2.72E+03	**-2.74E+03**
dataset 5	4.91E-03	4.90E-03	**4.31E-03**	6.72E-03	8.77E-03	8.76E-03	**7.42E-03**	1.10E-02	-3.05E+03	-3.04E+03	-3.04E+03	**-3.06E+03**
dataset 6	2.44E-03	**2.33E-03**	2.44E-03	2.88E-03	3.11E-03	**2.92E-03**	3.10E-03	3.56E-03	-3.32E+03	-3.32E+03	-3.32E+03	**-3.33E+03**
dataset 7	2.96E-03	3.10E-03	**2.87E-03**	3.98E-03	3.59E-03	3.83E-03	**3.56E-03**	5.00E-03	-3.34E+03	-3.33E+03	-3.33E+03	**-3.35E+03**
dataset 8	3.99E-03	3.99E-03	**3.97E-03**	2.82E-01	4.76E-03	4.75E-03	**4.63E-03**	4.76E-01	-3.48E+03	-3.48E+03	-3.48E+03	**-3.54E+03**
dataset 9	3.51E-03	**2.42E-03**	3.57E-03	1.89E-01	4.07E-03	**2.86E-03**	4.14E-03	3.76E-01	-3.65E+03	-3.63E+03	-3.65E+03	**-3.73E+03**

Table B.14: MAE, RMSE and AIC for GARCH, NA1-GARCH, NA2-GARCH and EGARCH models for Aviva datasets.

Datasets	MAE				RMSE				AIC			
	GARCH	NA1-GARCH	NA2-GARCH	EGARCH	GARCH	NA1-GARCH	NA2-GARCH	EGARCH	GARCH	NA1-GARCH	NA2-GARCH	EGARCH
dataset 1	**1.48E-03**	1.61E-03	1.64E-03	3.20E-03	**1.90E-03**	2.06E-03	0.002001-0.000102i	4.57E-03	-3.94E+03	**-3.95E+03**	-3925.569145+6.283185i	-3.95E+03
dataset 2	6.35E-03	6.35E-03	6.33E-03	**3.20E-03**	9.29E-03	9.29E-03	9.26E-03	**4.57E-03**	**-3.93E+03**	-3.92E+03	-3.92E+03	-3.92E+03
dataset 3	2.19E-03	**2.19E-03**	2.19E-03	2.60E-03	**3.44E-03**	**3.44E-03**	**3.44E-03**	4.30E-03	-3.54E+03	-3.53E+03	-3.53E+03	**-3.55E+03**
dataset 4	3.55E-03	**2.98E-03**	3.38E-03	3.51E-03	4.99E-03	**3.78E-03**	4.27E-03	6.29E-03	-3.40E+03	-3.38E+03	-3.38E+03	**-3.40E+03**
dataset 5	2.12E-03	**2.01E-03**	2.11E-03	3.34E-03	2.86E-03	**2.58E-03**	2.84E-03	4.07E-03	-3.53E+03	-3.52E+03	-3.52E+03	**-3.55E+03**
dataset 6	1.95E-03	2.32E-03	**1.94E-03**	2.21E-03	2.42E-03	2.83E-03	**2.42E-03**	2.80E-03	-3.52E+03	-3.51E+03	-3.52E+03	**-3.53E+03**
dataset 7	2.52E-03	2.54E-03	**2.51E-03**	2.54E-03	2.75E-03	2.78E-03	**2.74E-03**	3.56E-03	-3.73E+03	-3.72E+03	-3.72E+03	**-3.75E+03**
dataset 8	2.82E-03	2.72E-03	**2.45E-03**	2.37E-01	3.58E-03	3.40E-03	**3.00E-03**	3.98E-01	-3.99E+03	-3.98E+03	-3.98E+03	**-4.03E+03**
dataset 9	3.66E-03	3.86E-03	**3.54E-03**	5.25E-03	4.99E-03	5.17E-03	**4.92E-03**	6.77E-03	-4.05E+03	-4.05E+03	-4.04E+03	**-4.07E+03**

Table B.15: MAE, RMSE and AIC for GARCH, NA1-GARCH, NA2-GARCH and EGARCH models for BP datasets

Datasets	MAE				RMSE				AIC			
	GARCH	**NA1-GARCH**	**NA2-GARCH**	**EGARCH**	**GARCH**	**NA1-GARCH**	**NA2-GARCH**	**EGARCH**	**GARCH**	**NA1-GARCH**	**NA2-GARCH**	**EGARCH**
dataset 1	2.34E-03	**1.89E-03**	2.19E-03	2.52E-03	2.91E-03	**2.32E-03**	2.68E-03	2.90E-03	**-4.04E+03**	-4.03E+03	-4.03E+03	-4.04E+03
dataset 2	8.05E-03	8.05E-03	7.61E-03	**2.00E-03**	1.12E-02	1.12E-02	1.02E-02	**3.10E-03**	-4.02E+03	-4.01E+03	-4.00E+03	**-4.02E+03**
dataset 3	2.16E-03	**1.88E-03**	2.17E-03	2.00E-03	2.58E-03	**2.21E-03**	2.46E-03	3.10E-03	**-3.66E+03**	-3.65E+03	-3.64E+03	-3.64E+03
dataset 4	2.21E-03	2.19E-03	**1.99E-03**	6.89E-02	3.59E-03	3.69E-03	**3.05E-03**	9.49E-02	**-3.57E+03**	-3.56E+03	-3.56E+03	-3.55E+03
dataset 5	2.45E-03	2.45E-03	**2.45E-03**	1.32E-01	2.82E-03	2.82E-03	**2.81E-03**	2.21E-01	-3.83E+03	-3.82E+03	-3.82E+03	**-3.87E+03**
dataset 6	1.77E-03	1.77E-03	**1.76E-03**	4.51E-02	2.13E-03	2.13E-03	**2.09E-03**	7.48E-02	-3.93E+03	-3.93E+03	-3.92E+03	**-3.94E+03**
dataset 7	1.97E-03	1.96E-03	**1.80E-03**	2.05E-03	2.25E-03	2.21E-03	**2.04E-03**	2.50E-03	**-4.05E+03**	-4.05E+03	-4.04E+03	-4.04E+03
dataset 8	1.98E-03	2.02E-03	**1.95E-03**	3.60E-03	2.38E-03	2.60E-03	**2.35E-03**	4.67E-03	-4.18E+03	-4.17E+03	-4.17E+03	**-4.18E+03**
dataset 9	2.74E-03	**2.68E-03**	2.74E-03	5.57E-03	3.47E-03	**3.39E-03**	3.46E-03	6.67E-03	**-4.09E+03**	-4.09E+03	-4.09E+03	-4.09E+03

Table B.16: MAE, RMSE and AIC for GARCH, NA1-GARCH, NA2-GARCH and EGARCH models for GlaxoSmithKline datasets

Datasets	MAE				RMSE				AIC			
	GARCH	**NA1-GARCH**	**NA2-GARCH**	**EGARCH**	**GARCH**	**NA1-GARCH**	**NA2-GARCH**	**EGARCH**	**GARCH**	**NA1-GARCH**	**NA2-GARCH**	**EGARCH**
dataset 1	3.59E-03	**2.52E-03**	2.85E-03	4.28E-03	4.54E-03	**3.06E-03**	3.51E-03	5.35E-03	**-4.05E+03**	-4.04E+03	-4.04E+03	-4.05E+03
dataset 2	5.61E-03	1.06E-02	**3.78E-03**	5.08E-03	7.77E-03	1.30E-02	**5.40E-03**	6.41E-03	**-3.90E+03**	-3.89E+03	-3.89E+03	-3.89E+03
dataset 3	**8.63E-03**	9.94E-03	9.65E-03	6.74E-02	1.51E-02	1.61E-02	**1.48E-02**	7.72E-02	-3.22E+03	-3.21E+03	-3.21E+03	**-3.24E+03**
dataset 4	**2.88E-03**	4.09E-03	3.12E-03	6.74E-02	**4.05E-03**	5.20E-03	4.16E-03	7.72E-02	-2.48E+03	-2.49E+03	-2.47E+03	**-2.49E+03**
dataset 5	2.29E-02	**4.81E-03**	6.52E-03	1.44E-02	4.74E-02	**7.88E-03**	1.15E-02	3.21E-02	**-2.76E+03**	-2.73E+03	-2.72E+03	-2.75E+03
dataset 6	3.86E-03	3.86E-03	**3.52E-03**	4.42E-03	5.60E-03	5.59E-03	5.72E-03	**5.37E-03**	-3.03E+03	-3.02E+03	-2.96E+03	**-3.04E+03**
dataset 7	4.05E-03	2.74E-03	1.88E-03	**1.85E-03**	4.42E-03	3.07E-03	**2.20E-03**	2.42E-03	**-3.06E+03**	-3.05E+03	-3.05E+03	-3.06E+03
dataset 8	1.23E-02	**3.19E-03**	1.17E-02	5.35E-01	1.30E-02	**5.19E-03**	1.25E-02	8.86E-01	-3.40E+03	-3.39E+03	-3.40E+03	**-3.43E+03**
dataset 9	5.28E-03	2.72E-03	**2.60E-03**	4.49E-01	6.41E-03	3.30E-03	**3.12E-03**	7.83E-01	-3.71E+03	-3.69E+03	-3.69E+03	**-3.73E+03**

Table B.17: MAE, RMSE and AIC for GARCH, NA1-GARCH, NA2-GARCH and EGARCH models for Lloyds Bank datasets

Datasets	MAE				RMSE				AIC			
	GARCH	NA1-GARCH	NA2-GARCH	EGARCH	GARCH	NA1-GARCH	NA2-GARCH	EGARCH	GARCH	NA1-GARCH	NA2-GARCH	EGARCH
dataset 1	2.35E-03	2.57E-03	**2.27E-03**	9.43E-01	3.12E-03	3.37E-03	**2.98E-03**	9.47E-01	**-3.82E+03**	-3.77E+03	-3.79E+03	-4.31E+01
dataset 2	4.69E-03	3.67E-03	**3.51E-03**	3.56E-03	6.59E-03	5.41E-03	5.05E-03	**4.52E-03**	**-3.61E+03**	-3.60E+03	-3.60E+03	-3.60E+03
dataset 3	5.89E-03	**2.43E-03**	2.45E-03	1.74E-01	6.75E-03	**2.95E-03**	2.99E-03	1.78E-01	**-3.33E+03**	-3.32E+03	-3.31E+03	-1.17E+03
dataset 4	2.25E-03	2.27E-03	2.25E-03	**1.79E-03**	3.26E-03	3.29E-03	3.25E-03	**2.62E-03**	-3.29E+03	-3.28E+03	-3.28E+03	**-3.30E+03**
dataset 5	2.80E-03	2.47E-03	**1.78E-03**	2.19E-03	4.13E-03	3.43E-03	**2.22E-03**	2.93E-03	**-3.72E+03**	-3.72E+03	-3.71E+03	-3.71E+03
dataset 6	1.96E-03	**1.80E-03**	2.01E-03	2.00E-03	2.20E-03	**2.11E-03**	2.24E-03	2.37E-03	**-3.85E+03**	-3.84E+03	-3.84E+03	-3.84E+03
dataset 7	2.52E-03	**2.49E-03**	2.51E-03	4.54E-02	3.40E-03	**3.33E-03**	3.38E-03	7.03E-02	-3.93E+03	-3.93E+03	-3.93E+03	**-3.95E+03**
dataset 8	3.00E-03	3.00E-03	**2.96E-03**	2.63E-02	3.70E-03	3.71E-03	**3.69E-03**	1.22E-01	**-3.92E+03**	-3.91E+03	-3.91E+03	-3.71E+03
dataset 9	2.47E-03	**2.37E-03**	2.44E-03	1.39E-02	3.13E-03	**3.00E-03**	3.09E-03	1.97E-02	-3.75E+03	-3.74E+03	-3.74E+03	**-3.76E+03**

Table B.18: MAE, RMSE and AIC for GARCH, NA1-GARCH, NA2-GARCH and EGARCH models for Vodafone datasets

Datasets	MAE				RMSE				AIC			
	GARCH	NA1-GARCH	NA2-GARCH	EGARCH	GARCH	NA1-GARCH	NA2-GARCH	EGARCH	GARCH	NA1-GARCH	NA2-GARCH	EGARCH
dataset 1	1.68E-03	1.68E-03	**1.60E-03**	2.19E-03	2.08E-03	2.08E-03	**2.00E-03**	2.68E-03	-3.91E+03	-3.90E+03	-3.90E+03	**-3.92E+03**
dataset 2	7.69E-03	**7.61E-03**	7.68E-03	1.32E-02	1.19E-02	**1.18E-02**	1.19E-02	2.02E-02	-3.80E+03	-3.80E+03	-3.79E+03	**-3.81E+03**
dataset 3	5.06E-03	5.06E-03	**4.58E-03**	6.08E-03	9.46E-03	9.45E-03	**9.41E-03**	1.08E-02	-3.34E+03	-3.33E+03	-3.33E+03	**-3.34E+03**
dataset 4	5.14E-03	5.14E-03	**4.52E-03**	4.75E-03	5.42E-03	5.42E-03	**4.74E-03**	5.61E-03	-3.02E+03	-3.01E+03	-3.01E+03	**-3.04E+03**
dataset 5	3.93E-03	**3.93E-03**	3.93E-03	4.37E-03	6.67E-03	6.65E-03	6.67E-03	**6.08E-03**	-3.45E+03	-3.45E+03	-3.44E+03	**-3.46E+03**
dataset 6	2.70E-03	**2.69E-03**	2.98E-03	3.02E-03	3.94E-03	3.94E-03	**3.74E-03**	4.40E-03	-3.55E+03	-3.54E+03	-3.53E+03	**-3.56E+03**
dataset 7	**2.72E-03**	2.73E-03	**2.72E-03**	2.73E-03	3.21E-03	3.22E-03	**3.21E-03**	3.31E-03	-3.46E+03	-3.45E+03	-3.45E+03	**-3.47E+03**
dataset 8	2.78E-03	2.78E-03	**2.77E-03**	3.50E-03	4.58E-03	4.58E-03	**4.57E-03**	5.42E-03	-3.76E+03	-3.75E+03	-3.75E+03	**-3.78E+03**
dataset 9	**2.13E-03**	2.14E-03	2.22E-03	2.16E-03	**2.67E-03**	2.69E-03	2.79E-03	2.70E-03	-3.94E+03	-3.94E+03	-3.94E+03	**-3.95E+03**

Table B.19: MAE, RMSE and AIC for GARCH, NA1-GARCH, NA2-GARCH and EGARCH models for Allianz datasets

Datasets	MAE				RMSE				AIC			
	GARCH	**NA1-GARCH**	**NA2-GARCH**	**EGARCH**	**GARCH**	**NA1-GARCH**	**NA2-GARCH**	**EGARCH**	**GARCH**	**NA1-GARCH**	**NA2-GARCH**	**EGARCH**
dataset 1	3.47E-03	**3.40E-03**	3.60E-03	4.26E-03	5.07E-03	**4.95E-03**	5.26E-03	6.03E-03	-3.87E+03	**-3.88E+03**	-3.87E+03	-3.85E+03
dataset 2	7.66E-03	5.98E-03	7.96E-03	**5.93E-03**	1.06E-02	8.50E-03	1.10E-02	**6.92E-03**	**-3.70E+03**	-3.70E+03	-3.69E+03	-3.70E+03
dataset 3	8.15E-03	6.16E-03	**6.11E-03**	4.59E-01	1.51E-02	1.37E-02	**1.37E-02**	7.97E-01	-3.23E+03	-3.23E+03	-3.22E+03	**-3.28E+03**
dataset 4	3.68E-03	3.40E-03	**3.26E-03**	4.59E-01	4.03E-03	3.92E-03	**3.59E-03**	7.97E-01	-3.07E+03	-3.05E+03	-3.06E+03	**-3.11E+03**
dataset 5	1.37E-02	**7.98E-03**	9.87E-03	1.31E+00	2.04E-02	**1.36E-02**	1.67E-02	2.07E+00	-3.48E+03	-3.47E+03	-3.47E+03	**-3.49E+03**
dataset 6	1.86E-03	**1.70E-03**	1.81E-03	2.34E-03	2.29E-03	**2.09E-03**	2.23E-03	2.81E-03	-3.74E+03	-3.73E+03	-3.73E+03	**-3.74E+03**
dataset 7	2.33E-03	2.33E-03	**2.33E-03**	2.34E-03	3.14E-03	3.14E-03	3.14E-03	**2.81E-03**	-3.85E+03	-3.84E+03	-3.84E+03	**-3.85E+03**
dataset 8	2.83E-03	2.96E-03	**2.73E-03**	2.16E-02	3.39E-03	3.51E-03	**3.25E-03**	2.56E-02	-3.79E+03	-3.78E+03	-3.78E+03	**-3.79E+03**
dataset 9	2.70E-03	2.68E-03	**2.67E-03**	2.16E-02	3.44E-03	3.45E-03	**3.41E-03**	2.56E-02	-3.85E+03	-3.84E+03	-3.84E+03	**-3.89E+03**

Table B.20: MAE, RMSE and AIC for GARCH, NA1-GARCH, NA2-GARCH and EGARCH models for Anheuser-Busch datasets

Datasets	MAE				RMSE				AIC			
	GARCH	**NA1-GARCH**	**NA2-GARCH**	**EGARCH**	**GARCH**	**NA1-GARCH**	**NA2-GARCH**	**EGARCH**	**GARCH**	**NA1-GARCH**	**NA2-GARCH**	**EGARCH**
dataset 1	**2.76E-03**	2.86E-03	2.77E-03	3.62E-03	**3.61E-03**	3.72E-03	3.63E-03	5.00E-03	**-4.14E+03**	-4.12E+03	-4.12E+03	-4.12E+03
dataset 2	4.79E-03	4.71E-03	4.76E-03	**3.62E-03**	6.82E-03	6.71E-03	6.77E-03	**5.00E-03**	**-3.94E+03**	-3.93E+03	-3.93E+03	-3.92E+03
dataset 3	4.84E-03	4.84E-03	**4.48E-03**	5.02E-03	7.99E-03	7.99E-03	**7.62E-03**	8.22E-03	-3.46E+03	-3.45E+03	-3.45E+03	**-3.48E+03**
dataset 4	4.19E-03	4.16E-03	**3.69E-03**	4.83E-03	6.26E-03	6.22E-03	**5.43E-03**	6.30E-03	-3.10E+03	-3.10E+03	-3.10E+03	**-3.14E+03**
dataset 5	3.65E-03	3.64E-03	**3.44E-03**	4.60E-03	4.47E-03	4.45E-03	**4.34E-03**	5.87E-03	-3.15E+03	-3.14E+03	-3.14E+03	**-3.19E+03**
dataset 6	3.34E-03	3.34E-03	**2.94E-03**	1.07E+05	4.40E-03	4.40E-03	**3.88E-03**	2.74E+05	-3.21E+03	-3.20E+03	-3.20E+03	**-3.26E+03**
dataset 7	3.95E-03	**3.79E-03**	3.94E-03	3.61E-01	4.61E-03	**4.45E-03**	4.60E-03	6.49E-01	-3.23E+03	-3.22E+03	-3.22E+03	**-3.28E+03**
dataset 8	2.94E-03	**2.80E-03**	2.92E-03	7.73E-02	3.69E-03	**3.54E-03**	3.69E-03	1.03E-01	-3.40E+03	-3.39E+03	-3.39E+03	**-3.42E+03**
dataset 9	3.63E-03	**3.54E-03**	3.63E-03	4.39E-03	5.50E-03	**5.20E-03**	5.50E-03	5.78E-03	**-3.71E+03**	-3.70E+03	-3.70E+03	-3.70E+03

Table B.21: MAE, RMSE and AIC for GARCH, NA1-GARCH, NA2-GARCH and EGARCH models for Banco Santander datasets

Datasets	MAE				RMSE				AIC			
	GARCH	**NA1-GARCH**	**NA2-GARCH**	**EGARCH**	**GARCH**	**NA1-GARCH**	**NA2-GARCH**	**EGARCH**	**GARCH**	**NA1-GARCH**	**NA2-GARCH**	**EGARCH**
dataset 1	1.26E-03	1.56E-03	**1.26E-03**	4.12E-03	1.67E-03	2.08E-03	**1.66E-03**	6.32E-03	**-3.84E+03**	-3.83E+03	-3.84E+03	-6.30E+01
dataset 2	6.92E-03	6.92E-03	**3.70E-03**	4.12E-03	1.07E-02	1.07E-02	**5.74E-03**	6.32E-03	-3.73E+03	-3.73E+03	-3.72E+03	**-3.74E+03**
dataset 3	3.54E-03	**3.48E-03**	3.54E-03	3.79E-03	6.02E-03	**5.82E-03**	6.03E-03	6.35E-03	-3.45E+03	-3.45E+03	-3.44E+03	**-3.45E+03**
dataset 4	3.28E-03	3.42E-03	**3.28E-03**	3.43E-03	3.76E-03	3.87E-03	**3.70E-03**	4.10E-03	-3.31E+03	-3.30E+03	-3.30E+03	**-3.32E+03**
dataset 5	3.40E-03	3.97E-03	**3.31E-03**	3.57E-03	4.24E-03	4.96E-03	**4.22E-03**	4.35E-03	-3.56E+03	-3.54E+03	-3.53E+03	**-3.56E+03**
dataset 6	2.28E-03	2.03E-03	**1.98E-03**	2.38E-03	3.10E-03	2.99E-03	**2.71E-03**	3.33E-03	**-3.59E+03**	-3.59E+03	-3.58E+03	-3.59E+03
dataset 7	2.48E-03	**2.44E-03**	2.48E-03	1.60E-01	2.95E-03	**2.91E-03**	2.95E-03	3.09E-01	-3.53E+03	-3.52E+03	-3.52E+03	**-3.56E+03**
dataset 8	3.83E-03	**2.48E-03**	2.84E-03	1.22E+09	4.68E-03	**3.39E-03**	3.91E-03	9.10E+09	-3.72E+03	-3.71E+03	-3.71E+03	**-3.74E+03**
dataset 9	3.55E-03	3.91E-03	**3.54E-03**	3.89E-03	4.29E-03	4.67E-03	**4.28E-03**	4.73E-03	-3.78E+03	-3.77E+03	-3.77E+03	**-3.79E+03**

Table B.22: MAE, RMSE and AIC for GARCH, NA1-GARCH, NA2-GARCH and EGARCH models for Bayer datasets

Datasets	MAE				RMSE				AIC			
	GARCH	NA1-GARCH	NA2-GARCH	EGARCH	GARCH	NA1-GARCH	NA2-GARCH	EGARCH	GARCH	NA1-GARCH	NA2-GARCH	EGARCH
dataset 1	2.31E-03	**1.69E-03**	1.86E-03	2.45E-03	2.84E-03	**2.08E-03**	2.34E-03	2.94E-03	-3.92E+03	-3.90E+03	-3.90E+03	**-3.92E+03**
dataset 2	6.81E-03	**4.35E-03**	5.65E-03	9.32E-03	1.04E-02	**6.02E-03**	8.41E-03	1.31E-02	**-3.81E+03**	-3.77E+03	-3.80E+03	-3.81E+03
dataset 3	7.06E-03	7.37E-03	**5.91E-03**	8.65E-03	1.17E-02	1.19E-02	**1.12E-02**	1.36E-02	-3.33E+03	-3.32E+03	-3.31E+03	**-3.35E+03**
dataset 4	**2.69E-03**	3.58E-03	3.10E-03	3.74E-03	**3.91E-03**	4.52E-03	4.02E-03	4.56E-03	-2.80E+03	-2.82E+03	-2.81E+03	**-2.85E+03**
dataset 5	5.51E-03	2.83E-02	**5.29E-03**	6.67E-03	8.43E-03	3.22E-02	**7.69E-03**	1.07E-02	-3.06E+03	-3.07E+03	-3.06E+03	**-3.07E+03**
dataset 6	2.62E-03	2.86E-03	**2.54E-03**	3.15E-03	4.01E-03	3.89E-03	**3.41E-03**	4.37E-03	-3.31E+03	-3.31E+03	-3.31E+03	**-3.32E+03**
dataset 7	3.11E-03	3.11E-03	2.44E-03	**2.38E-03**	3.99E-03	3.99E-03	3.76E-03	**2.89E-03**	-3.19E+03	-3.19E+03	-3.17E+03	**-3.19E+03**
dataset 8	3.20E-03	**3.03E-03**	3.11E-03	1.44E+00	5.64E-03	**5.56E-03**	**5.56E-03**	2.74E+00	-3.37E+03	-3.36E+03	-3.36E+03	**-3.39E+03**
dataset 9	2.05E-03	2.82E-03	**1.92E-03**	2.45E-03	2.73E-03	3.79E-03	**2.65E-03**	2.95E-03	**-3.67E+03**	-3.66E+03	-3.63E+03	-3.65E+03

Table B.23: MAE, RMSE and AIC for GARCH, NA1-GARCH, NA2-GARCH and EGARCH models for Deutsche Bank datasets

Datasets	MAE				RMSE				AIC			
	GARCH	**NA1-GARCH**	**NA2-GARCH**	**EGARCH**	**GARCH**	**NA1-GARCH**	**NA2-GARCH**	**EGARCH**	**GARCH**	**NA1-GARCH**	**NA2-GARCH**	**EGARCH**
dataset 1	2.43E-03	**2.37E-03**	2.42E-03	3.90E-03	3.17E-03	**3.11E-03**	3.15E-03	5.13E-03	-4.02E+03	**-4.03E+03**	-4.03E+03	-4.00E+03
dataset 2	**4.50E-03**	**4.50E-03**	4.50E-03	1.15E-02	7.28E-03	**7.28E-03**	7.29E-03	1.89E-02	**-3.94E+03**	-3.93E+03	-3.93E+03	-3.90E+03
dataset 3	3.55E-03	4.16E-03	**3.55E-03**	6.80E-03	**7.13E-03**	8.82E-03	7.78E-03	1.06E-02	**-3.57E+03**	-3.52E+03	-3.56E+03	-3.53E+03
dataset 4	2.39E-03	2.63E-03	**2.36E-03**	4.58E-03	2.81E-03	3.48E-03	**2.74E-03**	5.14E-03	**-3.38E+03**	-3.37E+03	-3.37E+03	-3.37E+03
dataset 5	4.11E-03	**2.25E-03**	3.31E-03	4.56E-03	6.04E-03	**3.09E-03**	4.76E-03	5.30E-03	-3.63E+03	-3.62E+03	-3.62E+03	**-3.63E+03**
dataset 6	1.81E-03	1.81E-03	**1.73E-03**	2.56E-03	2.19E-03	2.19E-03	**2.09E-03**	3.08E-03	-3.74E+03	-3.74E+03	-3.74E+03	**-3.75E+03**
dataset 7	2.50E-03	2.55E-03	2.53E-03	**1.99E-03**	2.93E-03	2.98E-03	2.96E-03	**2.38E-03**	-3.76E+03	-3.76E+03	-3.76E+03	**-3.78E+03**
dataset 8	**2.34E-03**	**2.34E-03**	**2.34E-03**	4.18E-03	**3.38E-03**	**3.38E-03**	**3.38E-03**	5.26E-03	-3.93E+03	-3.92E+03	-3.92E+03	**-3.95E+03**
dataset 9	2.69E-03	2.69E-03	**2.67E-03**	5.36E-03	3.90E-03	3.90E-03	**3.88E-03**	6.95E-03	**-3.95E+03**	-3.94E+03	-3.94E+03	-3.93E+03

Table B.24: MAE, RMSE and AIC for GARCH, NA1-GARCH, NA2-GARCH and EGARCH models for Total datasets

Appendix C

Chapter 5

Tables of crude oil export and import countries

	Countries	Codes	Amounts in USD	Percentage
1	Saudi Arabia:	SA	$133.3	17%
2	Russia:	RU	$86.2	11%
3	Iraq:	IQ	$52.2	6.60%
4	United Arab Emirates:	AE	$51.2	6.50%
5	Canada:	CA	$50.2	6.40%
6	Nigeria:	NG	$38	4.80%
7	Kuwait:	KW	$34.1	4.30%
8	Angola:	AO	$32.6	4.10%
9	Venezuela:	VE	$27.8	3.50%
10	Kazakhstan:	KZ	$26.2	3.30%
11	Norway:	NO	$25.7	3.30%
12	Iran:	IR	$20.5	2.60%
13	Mexico:	MX	$18.8	2.40%
14	Oman:	OM	$17.4	2.20%
15	United Kingdom:	UK	$16	2%

Table C.1: Major exporting countries of crude oil and their proportions in global contributions(billion USD).

	Countries	Codes	Amounts in USD	Percentage
1	China:	CN	\$134.3	16.70%
2	United States:	US	\$132.6	16.50%
3	India:	IN	\$72.3	9%
4	South Korea:	KR	\$55.1	6.90%
5	Japan:	JP	\$45	5.60%
6	Germany:	DE	\$36.4	4.50%
7	Netherlands:	NL	\$35.4	4.40%
8	Spain:	ES	\$24.8	3.10%
9	Italy:	IT	\$23.7	3%
10	France:	FR	\$22.9	2.80%
11	Thailand:	TH	\$19.5	2.40%
12	United Kingdom:	UK	\$18.4	2.30%
13	Singapore:	SG	\$18.2	2.30%
14	Taiwan:	TW	\$16.1	2%
15	Belgium:	BE	\$14.8	1.80%

Table C.2: Major importing countries of crude oil and their proportions in global contributions(billion USD).

Table of observations statistics

Groups	# of News	Positive	Mean	Max	Min	Neutral	Negative	Mean	Max	Min
civil-unrest	71043	296	0.18	0.32	0.02	6579	64168	-0.51	-0.02	-1
commodity-prices	287901	113845	0.44	0.81	0.2	3063	170993	-0.44	-0.09	-0.8
consumption	86707	34888	0.41	1	0.02	20789	31030	-0.4	-0.02	-1
domestic-product	110969	31301	0.6	1	0.02	41862	37806	-0.63	-0.02	-1
exploration	16040	11941	0.53	0.61	0.5	13	4086	-0.53	-0.49	-0.57
foreign-exchange	263576	123301	0.49	0.7	0.24	12278	127997	-0.5	-0.24	-0.68
industrial-accidents	181273	6844	0.49	0.86	0.18	0	174429	-0.47	-0.06	-1
interest-rates	64834	16226	0.61	1	0.02	26187	22421	-0.7	-0.17	-1
taxes	4277	1767	0.54	0.86	0.1	1897	613	-0.54	-0.06	-0.78
transportation	16481	56	0.48	0.48	0.48	0	16425	-0.5	-0.4	-0.94
natural-disasters	138892	499	0.47	0.52	0.31	0	138393	-0.55	-0.16	-1
production	40595	16321	0.53	1	0.02	5119	19155	-0.51	-0.02	-1
products-services	1203242	938695	0.48	1	0.02	118201	146346	-0.54	-0.02	-1
war-conflict	244175	16981	0.55	1	0.1	8292	218902	-0.62	-0.14	-1
Total	**2730005**	**1312961**				**244280**	**1172764**			

Table C.3: The statistics summary of news sentiment for each group

Table of parameter estimates

Models	μ	θ	κ	σ	Q	P	λ	c1	c2	c3	c4	ρ	γ	a	b
Model 1	1.82E-01	9.51E-03	4.04E-02	1.83E-01	5.29E-03	2.36E+02									
Model 1-M	4.28E-01	2.35E-02	9.97E-02	1.33E-01	3.15E-03	5.45E+02						1.16E-02	2.92E-03	8.52E-01	6.33E-01
Model 2	1.41E-01	2.30E-02	3.16E-02	5.28E-02	6.91E-03	4.82E+02	9.31E-03								
Model 2-M	4.55E-01	7.01E-02	1.08E-01	1.43E-01	3.48E-03	3.50E+02	5.43E-03					1.54E-03	2.90E-03	6.45E-01	2.79E-06
Model 3	3.83E-01	2.35E-09	8.76E-02	1.30E-01	3.12E-03	1.71E+01	-1.35E-03	6.53E-03	-6.06E-04	1.01E-02	-4.52E-03				
Model 3-M	3.87E-01	5.29E-03	9.19E-02	9.11E-02	2.83E-03	1.23E+01	3.91E-03	1.26E-02	-3.17E-02	-5.84E-03	3.70E-03	3.06E-02	3.85E-02	2.00E+00	2.11E-05

Table C.4: Parameters estimation values of the models.

Appendix D

Software

The software developed for this research is provided in a CD attached to this thesis. It is also available online as a zip file at https://www.dropbox.com/sh/xqripd9k9nadqk8/AABOiZO7SL7051hdulruSaEqa?dl=0. The CD has the following file structure:

- Folder **ChapterThree** contains MATLAB files for the experiments discussed in Chapter 3. To reproduce any of the results of this chapter, one can begin with the run_chapter3.m script. The necessary EXCEL files with data are in the same folder and are read through relevant MATLAB functions.

  run_chapter3.m
  Before running run_chapter3.m from the MATLAB Command Window, a user first has to manually configure the following settings by choosing YES or NO (in Caps Lock):

  1. **saveAllResults**: whether to save an overall of the outcome results or not.
  2. **printExcelFiles**: whether to save the errors static in .csv files or not.
  3. **printVaR95**: whether to show the backtesting outcomes of VaR at 95% confidence level on the MATLAB Command Window (console) or not.
  4. **printVaR99**: whether to show the backtesting outcomes of VaR at 99% confidence level on the MATLAB Command Window (console) or not.
  5. **printParameters**: whether to save the estimated parameters values in a .csv file or not.
  6. **showPlots**: whether to show plots of volatility after running the script or not
  7. **saveTextFiles**: whether to save final summary of the results in a text file

or not.

After setting the above parameters, the user has to set the following variables before running the script:

1. **insamplePoints**: the number of in-sample observations.
2. **outOfSamplePoints**: the number of out-of-sample observations.
3. **index**: choosing an index ‘FTSE100’ or ‘SP500’.
4. **distributionType**: choosing a distribution assumption ‘NORMAL’ or ‘T’(Student t-distributed).

Calling the script will load the market data and the news data for the time period considered. It will then calculate the log-returns and realised volatility of market data. The script will call other scripts to calibrate the models considered in Chapter 3, using the maximum likelihood method. It will be followed by calculating the MAE and RMSE of volatilities estimated by the models. Other scripts will be called to evaluate the goodness of fit to the data of the models by calculating the maximum log-likelihood of each model. Finally the script will call some other scripts to compute the Value-at-Risk, assessment of the VaR method, and backtesting of the models. The functions that will be called by this script are:

1. realisedVolatility.m: to calculate the realised volatility.
2. calibrateGARCH.m: to calibrate the GARCH model when the distribution is normal.
3. calibrateGARCH_t.m: to calibrate the GARCH-t model when the distribution is Student t-distributed.
4. calibrateNA1_GARCH.m: to calibrate the NA1-GARCH model when the distribution is normal.
5. calibrateNA1_GARCH_t.m: to calibrate the NA1-GARCH-t model when the distribution is Student t-distributed.
6. calibrateTGARCH.m: to calibrate the TGARCH model when the distribution is normal.
7. calibrateTGARCH_t.m: to calibrate the TGARCH-t model when the distribution is Student t-distributed.

8. calibrateEGARCH.m: to calibrate the EGARCH model when the distribution is normal.
9. calibrateEGARCH_t.m: to calibrate the EGARCH-t model when the distribution is Student t-distributed.
10. GARCH.m: to forecast one-step ahead volatility using GARCH model.
11. NA1_GARCH.m: to forecast one-step ahead volatility using NA1_GARCH model.
12. TGARCH.m: to forecast one-step ahead volatility using TGARCH model.
13. EGARCH.m: to forecast one-step ahead volatility using EGARCH model.
14. LLH.m: to calculate the log-likelihood of the models when the distribution is normal.
15. LLH_t.m: to calculate the log-likelihood of the models when the distribution is Student t-distributed.
16. AIC.m: to calculate the Akaike information criterion (AIC) of the models.
17. BIC.m: to calculate the Bayesian information criterion (BIC) of the models.
18. VaR.m: to calculate the Value-at-Risk.
19. kupiec.m: to backtest the models using the Kupiec test.
20. christofferns.m: to backtest the models using the christofferns test.

- Folder **ChapterFour** contains MATLAB files for the experiments discussed in Chapter 4. To reproduce any of the results of this chapter, one can begin with the run_chapter4.m script. The necessary EXCEL files with data are in the same folder and are read through relevant MATLAB functions.

run_chapter4.m
Before running run_chapter4.m from the MATLAB Command Window, a user first has to manually configure the following settings:

1. **index**: choosing an index 'FTSE100' or 'EUROSTOXX50'.
2. **newsSource**: choosing the news data source if there are more than one is available.

3. **decayMethod**: choosing the decay function ('Exponential' or 'Hill').
4. **printParameters**: whether to save the estimated parameters values in a .csv file or not, by choosing YES or NO (in Caps Lock).
5. **printerrors**: whether to save the errors static in a .csv file or not, by choosing YES or NO (in Caps Lock).
6. **printResults**: whether to save an overall of the outcome results in a .csv file or not, by choosing YES or NO (in Caps Lock).
7. **showPlots**: whether to show plots of volatility after running the script or not, by choosing YES or NO (in Caps Lock).
8. **RVstepSize**: the step size of realised volatility (how many previous days should be taken to calculate daily realised volatility).
9. **firstObservation**: choose where the first observation should be started from the excel file (starting date).
10. **lastObservation**: choose where the last observation should be ended from the excel file (ending date).
11. **insamplePoints**: the number of in-sample observations.
12. **outOfSamplePoints**: the number of out-of-sample observations.

After setting the above parameters, calling the script will load the market data and the news data for each asset over the time period considered. It will then calculate the log-returns and realised volatility of market data. The script will call other scripts to calibrate the models considered in Chapter 4, using the maximum likelihood method. It will be followed by computing the MAE and RMSE of volatilities estimated by the models. Finally the script will call some other scripts to evaluate the goodness of fit to the data of the models by calculating the maximum log-likelihood of each model. The functions that will be called by this script are:

1. realisedVolatility.m: to calculate the realised volatility.
2. settings.m: setting the initial values of the model parameters.
3. calibrateGARCH.m: to calibrate the GARCH model when the distribution is normal.

4. calibrateNA1_GARCH.m: to calibrate the NA1-GARCH model when the distribution is normal.
5. calibrateNA2_GARCH.m: to calibrate the NA2-GARCH model when the distribution is normal.
6. calibrateEGARCH.m: to calibrate the EGARCH model when the distribution is normal.
7. GARCH.m: to forecast one-step ahead volatility using GARCH model.
8. NA1_GARCH.m: to forecast one-step ahead volatility using NA1_GARCH model.
9. NA2_GARCH.m: to forecast one-step ahead volatility using NA2_GARCH model.
10. EGARCH.m: to forecast one-step ahead volatility using EGARCH model.
11. plotOneResult.m: to plot two variables (realised volatility + volatility estimated by a model).
12. plotTwoResults.m: to plot three variables (realised volatility + volatilities estimated by two different models).
13. LLH.m: to calculate the log-likelihood of the models when the distribution is normal.
14. AIC.m: to calculate the Akaike information criterion (AIC) of the models.
15. BIC.m: to calculate the Bayesian information criterion (BIC) of the models.

- Folder **ChapterFive** contains MATLAB files for the experiments discussed in Chapter 5. To reproduce any of the results of this chapter, one can begin with the run_chapter5.m script. The necessary EXCEL files with data are in the same folder and are read through relevant MATLAB functions.

run_chapter5.m

Before running run_chapter5.m from the MATLAB terminal, a user first has to manually configure the following settings:

1. **model**: choosing a model to calibrate. The models are: One factor model (OF), One factor model with risk premium (OFRP) and One factor model

with risk premium and seasonality (OFRPS). (ALL) is the option if the user wants to calibrate all the models together.

2. **saveResults**: whether to save the outcome results to .csv files or not, by choosing YES or NO (in Caps Lock).
3. **saveErrors**: whether to save the errors static to a .csv file or not, by choosing YES or NO (in Caps Lock).
4. **showPlots**: whether to show plots of spot and futures prices (real and estimated) after running the script or not, by choosing YES or NO (in Caps Lock).
5. **useScoresOf**: which scores should be used, 1impact' or 'sentiment'?
6. **operator**: how to incorporate news data to the models? 'additive' or 'multiplicative'?
7. **newsMethod**: which method should be used, 'method1':= add news as external regressors, or 'method1':= use the chapter 4 approach (the scaling factor (NA2)).
8. **impactMethod**: which impact method should be chosen, 'Mean', 'Greater-WeightToLastNews', 'Weighted' or 'Aggregated'?
9. **Precision**: number of function evaluations and maximum iterations.
10. **in_sample**: the number of in-sample observations.
11. **out_of_sample**: the number of out-of-sample observations.

After setting the above parameters, calling the script will load the crude oil spot and futures prices, and the macroeconomic news data for the time period considered. It will then perform a Kalman calibration to estimate the model parameters as explained in Chapter 5, by calling the main script (main.m). The main.m script will forecast the spot and futures prices followed by computing the MAE, RMSE, AIC and AICc of the models. The functions that will be called by run_chapter5.m and main.m scripts are:

1. plotCommodity.m: to plot the spot price (real and estimated).
2. plotFutures.m: to plot the futures prices (real and estimated).
3. importResultsToFiles.m: to import the final results (including the estimated parameters value, log-likelihood values, MAE, MRSE, AIC, AICc, etc.) to

a .csv file.

4. saveErrorsValues.m: to save MAE and RMSE in a .csv file.
5. models.m: to forecast the spot and futures prices, or to estimate the chosen model.
6. futuersErrors.m: to compute the MAE and MRSE of futures prices estimated by the models.
7. AIC.m: to calculate the Akaike information criterion (AIC) and the second order of Akaike information criterion (AICc) of the models.

Please note that all the script and the function files in the above folders have detailed comments regarding the actual code.