AN INVESTIGATION INTO USING NEWS ANALYTICS DATA IN GARCH TYPE VOLATILITY MODELS

Sergey P. Sidorov

A Thesis Submitted for the Degree of MPhil at
Brunel University,
the School of Information Systems, Computing and Mathematics

Contents

List of Figures

List of Tables

Abstract

In the work we study different dynamic volatility models. We consider the family of ARCH and GARCH models to compare the performance of the models using both unconditional coverage Kupiec's test and the test of conditional coverage proposed by Christoffersen. In-sample estimation procedure and out-of-sample evaluation will be based on General Electric stock market closing daily prices (January 2, 2008 - December 31, 2010).

We consider different volatility models augmented with news analytics data to examine the impact of news intensity on stock volatility. First we consider two types of GARCH models: augmented with volume and augmented with news intensity. Based on empirical evidences for some of FTSE100 companies it will be shown 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.

Then we compare GARCH model with jumps and GARCH–Jumps model augmented with news intensity using likelihood ratio test.

The study shows that the problem of examining the impact of news intensity on volatility is far more sophisticated than it might seem at first sight. Some hypothesists and suggestions for future work are proposed in the final chapter.

Acknowledgements

My profound thanks to my supervisor Dr Paresh Date with whom I have the great pleasure of working. In addition I would like to express my immense gratitude to Vladimir Balash for generous assistance, and Prof Gautam Mitra for creative support and for the kindly provided opportunity to use Raven Pack news analytics data.

This work could not have been written without financial support of Russian Government Programme "National Research University" and organizational support of Tatyana Zakharova, pro-vice chancellor of Saratov State University, as well as the staff of managerial board of the programme "National Research University".

I would like to express my gratitude to Prof Brendan McCabe and Keming Yu for helpful comments and remarks.

Chapter 1

Introduction

In this work we focus on the stock volatility and some adequate models of its dynamics. GARCH models are used by many researchers and practitioners to generate volatility forecasts. GARCH models give useful and quite reliable estimates of the conditional stock variance. Since the appearance, the ARCH and GARCH models proved their effectiveness. It is known that one of advantages of the ARCH and GARCH models is simple parameterization.

On the other hand, this models cannot capture the asymmetric effect discovered by Black (1976). This effect occurs when an unexpected drop in price (bad news) increases predictable volatility more than an unexpected increase in price (good news) of similar magnitude. This effect suggests that a symmetry constraint on the conditional variance function in past is inappropriate. So we are going to apply some asymmetric models (including EGARCH, GJR-GARCH and Threshold GARCH). To validate the calibrated model we will use out-of-sample evaluation procedure.

Since GARCH model cannot explain observed changes in stock volatility, the question is what are the core factors driving its behavior. In our quest we will rely on the Mixture of Distribution Hypothesis (MDH). MDH suggests that the stock volatility is closely related to the information frequency. We are going to analyze the impact of news on stock volatility by considering two types of models:

1. *Augmented GARCH models*. Augmented GARCH models allows us to

indirectly test the Mixture of Distribution Hypothesis (MDH). We use two different proxy for the mixing variable

- Following the study of Lamoureux and Lastrapes Lamoureax and Lastrapes [1990] we will suppose that trading volume can be considered as a proportional proxy for information arrivals to the market. We will show that once contemporaneous volume is included as an exogenous variable in the model, the GARCH persistence effect diminishes.
- Kalev and al. (2004), Cousin and Launois [2006] considered the "daily number of press releases on a stock" (news intensity) as the most appropriate explanatory variable in the basic equation of GARCH model. Following their studies we will examine the GARCH model with news intensity using news analytics data from Raven Pack, one of the biggest providers of news analytics in the world.

2. *GARCH–Jumps Model augmented with news analytics data*.

To calibrate this models we will use two types of input data: stock prices (source: Yahoo!Finance); news sentiment scores (source: RavenPack News Scores.

The work is organized as follows.

Chapter 2 provides a brief review of the literature about ARCH and GARCH models, the relation between information and conditional volatility.

Chapter 3 introduces notation and describes two methods of calibration of models: the maximum likelihood estimator and generalized moments method.

Chapter 4 presents a short introduction to volatility measurement.

In the chapter 5 one can find description of the family of ARCH and GARCH models. Asymmetric GARCH models also are presented. Empirical study consists of in-sample estimation and out-of-sample forecast evaluation of VaR to estimate predictive properties of the models.

Chapter 6 of the work presents augmented GARCH models. Empirical results are obtained based on two different data sets: stock prices and Raven

Pack news wires.

Chapter 7 presents two different GARCH models with jumps. We will consider the problem of calibration of the models and give some empirical results.

Chapter 8 concludes and provides some perspectives for future research.

Chapter 2

Literature Review

It is well-known that financial data are usually serial dependent. Moreover, distribution of the data is often heavy-tailed, asymmetric and therefore not Gaussian and volatility changes over time. The ARCH model was introduced by R. Engle (Engle [1982]). The model describes quite well the stylized facts of financial data and is also relatively simple and stationary. R. Engle called his model autoregressive conditionally heteroskedastic – ARCH, because the conditional variance (squared volatility) is not constant over time and shows autoregressive structure. This model is a convenient way of modelling time-dependent conditional variance. Some years later, Bollerslev [1986] generalized this model as the GARCH model (Generalized Autoregressive Conditional Heteroscedasticity).

Since the appearance, the ARCH and GARCH models proved their effectiveness. It is clear that one of advantages of the ARCH and GARCH models is simple parametrization. On the other hand, this models cannot capture the asymmetric effect discovered by Black in the paper Black [1976]. This feature was also confirmed by Kenneth R. French, G. William Schwert and Robert F. Stambaugh French et al. [1987], D. Nelson Nelson [1990], and G. William Schwert Schwert [1990], among others.

This effect occurs when an unexpected drop in price (bad news) increases predictable volatility more than an unexpected increase in price (good news) of similar magnitude. This effect suggests that a symmetry constraint on the conditional variance function in past innovations is inappropriate.

To capture the asymmetric effects a few models were introduced. The most commonly used asymmetric GARCH models are:

1. Exponential GARCH (EGARCH) model (Nelson [1991]);
2. The Quadratic GARCH (QGARCH) model (Sentana [1995]);
3. The Glosten-Jagannathan-Runkle GARCH (GJR-GARCH) model (Glosten et al. [1993]);
4. The Threshold GARCH (TGARCH) model (Zakoian [1994]).

There are well-known empirical studies of the positive contemporaneous correlation between trading volume and price volatility. In the paper Karpoff [1987] one can find the review of previous research on the relation between price changes and trading volume in financial markets. It cites about 20 papers that examine this relation in financial markets including equities, futures, currencies, and Treasury bills. Some of these papers also document an asymmetry in the relation; positive price shocks are associated with larger volumes than negative price shocks.

Trading volume is one of the most favored proxies for news arrivals. It can be explained by the following way: the more specific news arrives about a given stock (or company), the more investors will interpret the effects of that news differently, and thus the more investors will have an incentive to trade as their expectations about future returns diverge. There are several theoretical models which have been proposed to explain this relationship, for example, "asymmetric information" model, "differences in opinion" model, the sequential information arrival hypothesis, and the "mixture of distributions" model.

Mixture of Distribution Hypothesis (MDH) was proposed for the first time by Clark (Clark [1973]). It assumes that the joint distribution of daily return and volume can be modeled as a mixture of multivariate normal distributions.[1] The main idea of MDH is that returns on financial assets are generated from a mixture of distributions in which the stochastic mixture variable

[1] Although volume is a non-negative variable whereas a normal variable could take negative values, it can be assumed that volume is normal distributed. It follows from the facts that stocks of analyzed companies are usually highly liquid, means of volumes are large quantities (e.g. more than 10^7 for HSBC), and 6 sigma point is typically above zero.

is considered to be the rate of arrival of information flow into the market. Specifically, they are contemporaneously dependent on an underlying mixing variable that represents the flow of information. As a consequence, the variance of returns at a given interval is expected to be proportional to the rate of information arrival at the market. The development of MDH can be found in the papers Epps and Epps [1976], and Tauchen and Pitts [1983].

The paper Lamoureax and Lastrapes [1990] examine the validity of MDH for daily stock returns. It exploits the implication of the MDH that the volatility of daily price increments is positively related to the rate of daily information arrival. They suggest the daily trading volume as a measure of the amount of information flowing into the market every day. The authors used daily trading volume as a proxy for mixing variable and showed that ARCH and GARCH coefficients vanish if volume is included as an explanatory variable in the GARCH model.

Lamoureux and Lastrapes (Lamoureax and Lastrapes [1990]) used daily trading volume as a proxy for the mixing variable. They showed that the introduction of volume as an exogenous variable in the conditional variance equation eliminates the persistence of GARCH effects as measured by the sum of the GARCH parameters.

After the work of Lamoureax and Lastrapes [1990], a number of papers studying this issue have been appeared. However, the findings are not consistent. For example, the paper Sharma et al. [1996] tests for the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) effects in stock market indicator returns using the NYSE daily return and volume data for four years. The study shows that the market indicator returns are best described by the simple GARCH model (without volume as a mixing variable). Moreover, it was shown that the inclusion of volume as a proxy for information arrival in the conditional variance model does not necessary lead to the decrease of the GARCH effects.

In the paper Arago and Nieto [2005] investigate the issue for market index data for nine countries. It was found that volume effects do not cancel out GARCH effects at the country index level. In the paper Arago and Nieto [2005] unexpected trading volume uses as a proxy variable for the information flow. It was shown that the inclusion of trading volume does not reduce the persistence of conditional volatility.

On the other hand, there are some papers with findings analogous to Lamoureax and Lastrapes [1990]. For example, Pagunathan and Peker Ragunathan and Peker [1997] a strong contemporaneous effect of trading volume on volatility in the Sydney Futures Exchange. In the paper Miyakoshi [2002] one can find that the inclusion of the trading volume variable in EGARCH models eliminates the ARCH/GARCH effect for individual stocks as well as for the index on the Tokyo Stock Exchange. Bohl and Henke (2003) examine Polish stock markets and obtain similar results.

The paper Xiao et al. [2009] examines the quantitative relationship between volume and volatility in the Australian Stock Market. The authors study interaction of GARCH and volume effects on the entire available data for the Australian All Ordinaries Index. They showed that GARCH model testing and estimation is impacted by trading volume. In the paper the daily trading volume was used as a proxy for information arrival time. It is shown that the daily trading volume have significant explanatory power regarding the variance of daily returns. They also studied the impact of firm size on volatility. It was shown that the actively traded stocks with larger number of information arrivals per day have a larger impact of volume on the variance of daily returns, while low trading volume and small firm lead to a higher persistence of GARCH effects in the estimated models.

Thus, it is still interesting to investigate this issue with a new proxy variable for the information flow. One of them might be the news intensity (the number of news about a company at the day t is called the news intensity at the day t). There are not so much studies tried to examine relation between flows of information and stock (or market) volatility. One of the reason for that is the difficulty to find good empirical proxy of information arrivals.

Different measures of information arrivals were employed in variety of empirical studies in order to test the impact of the rate of information on the market volatility:

- macroeconomic news, Ederington and Lee [1993];
- the number of daily newspaper headlines and earnings announcements, Berry and Howe [1993];
- the number of specific stock market announcements, Mitchell and Mulherin [1994].

It was written in the paper Kalev et al. [2004] that ”the use of unconditional volatility measures such as absolute daily market returns in these studies often generates weak or inconclusive results regarding the news-volatility relation". Indeed, a news intensity is known to be quite noisy, and the presence of conditional heteroscedasticity in the returns time-series may significantly destroys the quality of the results.

It is worth mentioning the paper Andersen [1996] in which it was shown that different types of news have a different impact on the conditional stock volatility.

In the paper of Kalev et al. [2004] it was used firm-specific announcements as a proxy for information flows. It was shown that there exists a positive and significant impact of the arrival rate of the selected news variable on the conditional variance of stock returns on the Australian Stock Exchange in a GARCH framework. They split all their press releases into different categories according to their subject.

Some authors have worked specifically on the French market (Cousin and Launois [2006]).

Chapter 3

Preliminaries

3.1 Notation

Let X be a random variable defined on a probability space (Ω, Σ, P). Let F be the cumulative distribution function of X:

$$F(x) := P(X \leq x).$$

The mean of the random variable X (or the expected value of X) can be defined as

$$\mu = \mathbb{E}(X) = \int_{\mathbb{R}} x dF(x). \tag{3.1}$$

If the integral (3.1) exists, i.e. the random variable X has the mean μ, then the variance of X is defined by

$$\sigma^2 = Var(X) = \mathbb{E}((X - \mu)^2), \tag{3.2}$$

the variance is defined as the average squared deviation of each number from its mean.

Standard deviation of X is the square of the variance:

$$\sigma = \sqrt{Var(X)}.$$

The variance measures how spread out a distribution is. The variance is the one of the most used volatility measures.

Consider the case when X is a discrete random variable. Let $x_1, \ldots, x_N$ be a population of the length N. Then the population mean is defined as

$$\mu = \frac{1}{N} \sum_{i=1}^{N} x_i.$$

The population variance is defined by

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2.$$

When we are dealing with some empirical data, it is difficult or impossible to know exactly distribution of a random variable X. So we can not apply (3.1) to find the value of the variance of X. Then usually the variance is estimated based on a sample.

Let $y_1, \ldots, y_n$ be a sample (i.e. finite number of realizations of a discrete or continuous random variable X). Then the sample mean

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$$

is estimator of the mean of the random variable X.

To estimate the variance $Var(X)$ of the random variable X we can use two estimators.

1. The variance of the sample

   $$s_n^2 = \frac{1}{n} \sum_{i=1}^{n} (y_i - \bar{y})^2.$$

   Note that $\mathbb{E}(s_n^2) = \frac{n-1}{n}\sigma^2$.

2. The unbiased estimator of the variance

   $$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2.$$

Unbiased means that $\mathbb{E}(s^2) = \sigma^2$, i.e. the expected value of the sample estimator coincides with the theoretical value of the variance.

Consider a random variable X distributed normally with mean μ and variance σ^2, $X \sim \mathcal{N}(\mu, \sigma^2)$. About 68% of values drawn from a normal distribution are within one standard deviation s away from the mean; about 95% of the values lie within two standard deviations; and about 99.7% are within three standard deviations. This fact is known as the 68-95-99.7 rule, or the empirical rule, or the 3-sigma rule. To be more precise, the probability of that X lies between $\mu - n\sigma$ and $\mu + n\sigma$ is given by

$$P(\mu - n\sigma \leq X \leq \mu + n\sigma) = \Phi(n) - \Phi(-n), \tag{3.3}$$

where

$$\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp^{-t^2/2} dt.$$

3.2 Calibration of models

3.2.1 The maximum likelihood estimator

In this subsection we will shortly describe this method as well as its shortcomings and advantages. We can refer the reader to the paper Le Cam [1990] for a deep study and a related bibliography.

We will use maximum likelihood for calibration the GARCH model in Chapter 5.

Let

$$x_1, x_2, \ldots, x_n \tag{3.4}$$

be a vector of observations of length n. We suppose that x_i are realizations of a random variable ζ with an unknown distribution density f. It is assumed that the density $f(\cdot)$ belongs to a parameterized family of distributions

$$\mathcal{F} := \{f(\cdot|\theta) : \ \theta \in \Theta\}.$$

Denote θ^* the vector of the true values of parameters. We need to find θ^0 so that θ^0 would be close to θ^* as much as possible in some sense.

The joint density function for observation sequence (3.4) is

$$f(x_1, \dots, x_n|\theta) = f(x_1|\theta) \cdot f(x_2|\theta) \cdot \dots \cdot f(x_n|\theta) = \prod_{t=1}^{n} f(x_t|\theta).$$

Let us to view the likelihood function as a probability density for θ, and to think of $f(x_1, \dots, x_n|\theta)$ as the conditional density of θ given $x_1, \dots, x_n$:

$$L(\theta|x_1, \dots, x_n) = f(x_1, \dots, x_n|\theta).$$

The function $L(\theta|x_1, \dots, x_n)$ is called *the likelihood function*.

We need to find $\theta^0 \in \Theta$ that maximizes the likelihood function $L(\theta|x_1, \dots, x_n)$ over all $\theta \in \Theta$, given $x_1, \dots, x_n$. In many cases it is much easy to maximize the logarithm value of the function:

$$\log L(\theta|x_1, \dots, x_n) = \sum_{t=1}^{n} \log f(x_t|\theta). \tag{3.5}$$

The function (3.5) is called *the log-likelihood function*.

There are many models in which a maximum likelihood estimator can be found as an explicit function of the observed data $x_1, \dots, x_n$, i.e. in Gaussian case.

However, for the most models there is not closed-form solution to the maximization problem

$$\log L(\theta|x_1, \dots, x_n) \to \max_{\theta \in \Theta}, \tag{3.6}$$

and a maximum likelihood estimation has to be found numerically using optimization methods.

The maximum likelihood estimator θ^0 is defined by

$$\theta^0 = \arg\max_{\theta \in \Theta} \log L(\theta|x_1, \dots, x_n).$$

Thus we are choosing θ to maximize the probability of occurrence of the observation $x_1, \dots, x_n$.

We point out some attractive asymptotic properties of the maximum-likelihood estimator:

1. **Consistency:** the maximum likelihood estimator θ^0 converges in probability to the value θ^* being estimated, i.e. $\theta^0 \to \theta^*$ as the number of observations n tends to ∞.

2. **Asymptotic normality:** as the sample size increases, the distribution of the MLE tends to the Gaussian distribution with mean θ and covariance matrix equal to the inverse of the Fisher information matrix. (see e.g. Myung and Navarro [2005]).

3. **Efficiency**, i.e., it achieves the Cramér–Rao lower bound when the sample size tends to infinity. This means that no asymptotically unbiased estimator has lower asymptotic mean squared error than the MLE.

It is worth noting that in many cases problem (3.6) is non-convex and, thus, finding its exact solution is in principle a difficult task. Optimization methods may include gradient climbing algorithms such as Newton-Raphson and EM algorithm.

Moreover, local minima are indeed a problem: different initial sets of parameters yield a different local minimum. But the problem is not only multiple local minima, also that quite often solutions are in the boundary of the feasible parameter space, e.g. some variances of the noises are estimated to have value zero. We will encounter this problem in Chapter 7.

3.2.2 Generalized moments method

Sometimes we do not know the shape of the distribution function of the data, but we can assume which of parameterized set of functions our distribution belongs to. In such cases the maximum likelihood estimation can not be apply, but the generalized moment method can be useful.

Let $y_1, y_2, \dots, y_n$ be a sample from a distribution governed by parameter θ. A function $e(y)$ is said to be *sample statistic* if it is a function of sample observations $y_1, y_2, \dots, y_n$ alone.

We say that an estimator $e(y)$ is *unbiased* if $\mathbb{E}(e(y)) = \theta$. If $\mathbb{E}(e(y)) \neq \theta$ then the value $\mathbb{E}(e(y)) - \theta$ is the bias of the estimator $e(y)$.

An estimator $e(y)$ is said to be *minimum variance unbiased* if it is unbiased and for every any other unbiased estimator e^* of θ we have $\text{Var}(e(y)) \leq \text{Var}(e^*(y))$.

We say that an estimator $e(y)$ is consistent if $e(y)$ converges to θ in probability as $n \to \infty$.

First we consider classical moment methods. Classical moment methods is a simple method for estimating unknown parameters in different statistical models.

Let $\theta \in \mathbb{R}^m$ be a vector of parameters (moments) that characterize the distribution of random variable y. Let the distribution function of the random variable y belongs to parameterized family of distributions $\{F(\cdot, \theta) : \theta \in \Theta\}$. The k-th moment (if it exists) of distribution of random variable y is defined as

$$m_k(\theta) = \mathbb{E}(y^k) = \int_{\mathbb{R}} y^k dF(y, \theta). \tag{3.7}$$

If we have sample observations $y_1, y_2, \ldots, y_n$ (considered as n independent random variables), then we can find k-th sample moment as follows

$$\hat{m}_k = \frac{1}{n} \sum_{i=1}^{n} y_i^k \tag{3.8}$$

Suppose we know the values some of the moments $\hat{m}_k$, $k \in I$, where $I \subset \mathbb{N}$, $|I| = k$. The main idea of method of moments is to estimate unknown parameters θ by matching "theoretical" moments (3.7) and sample moments (3.8) of the same orders:

$$m_k(\theta) = \hat{m}_k, \; k \in I. \tag{3.9}$$

In general, method of moments estimators are consistent whenever the Law of Large Numbers ensures that the sample moments in the data-generating process converge in probability to the corresponding population moments.

The classical method of moments can be applied if $m = k := |I|$. It follows from the fact that the number of unknown parameters is equal to the number of equations in (3.9) that the system (3.9) has a solution. If $m < k$ then (3.9) does not have a solution.

The first idea is to exclude $k - m$ equations and then apply the classical method of moments. The question is which of equations must be discard? The generalized method of moments technique uses all k moment conditions by weighting them, i.e. it chooses an estimator that balances each moment condition against the others. A GMM estimator may satisfy no one moment condition, but it may come close to satisfying them all.

Let data be a finite number of realizations of the process x_t, $t = 1, 2, \ldots, N$. Let the model satisfy the moment conditions:

$$\mathbb{E}(f(x_t, \theta^*)) = 0 \tag{3.10}$$

where $x_t \in \mathbb{R}^r$ is an vector of observable variables, $\theta^* \in \mathbb{R}^m$ is a vector of true value of parameters and $f : \mathbb{R}^r \times \mathbb{R}^m \to \mathbb{R}^k$ is a vector valued function. We assume that on the parameter space $\mathbb{E}(f(x_t, \theta)) = 0$ if and only if $\theta = \theta^*$.

Sample analogue of (3.10) can be written as

$$g_N(\theta) = \frac{1}{N} \sum_{t=1}^{N} f(x_t, \theta).$$

Suppose that we have a sequence of $k \times k$ positive semi definite matrix W_N converging to a positive definite matrix W_0. Then, GMM estimator is defined as

$$\hat{\theta} = \arg \inf_{\theta} (g'_N(\theta)\, W_N\, g_N(\theta)).$$

GMM does not guarantee an efficient estimator, but it does provide a consistent estimator, and its weighting scheme is more efficient than the simpler unweighted scheme Ravi et al. [2002].

We refer readers to the paper Hall [2004] for more details.

Chapter 4

Volatility

4.1 What is volatility

Volatility can be described as the relative rate at which the price of a security moves up and down. Volatility may be found by calculating the annualized standard deviation of daily change in price. Usually volatility describes the behavior of a financial instrument for a specified period of time, i.e. 1 day or 30 days or 90 days.

The measurement of the volatility of a financial instrument may be based on historical prices over the specified period. In this case it is obvious that the last observations the most recent price are needed to get more precise prediction of this behavior. It is supposed that if the price of a stock moves up and down rapidly over short time periods, then it has high volatility. And vice-versa: if the price has small changes or almost never changes, it has low volatility.

If we consider the problem option pricing then there are several variables that are of interest in financial engineering. They include

- the current asset price,
- the strike price,
- time to maturity,

- the risk free rate,
- volatility.

The first four of them are known. Their values can be directly obtained or derived from current market data. The volatility of a stock is the only variable that can not be so easily found.

4.2 Volatility measurement

4.2.1 Variance

It is supposed that volatility is a measure of the range (dispersion) of an asset price about its mean value over a certain amount of time. Then it follows that volatility is connected to the variance of an asset price. A stock is said to be volatile if the price will vary greatly over time. Conversely, a less volatile stock will have a price that will deviate relatively little over time.

The variance measures how spread out a distribution is. The variance is the one of the most used volatility measures.

4.2.2 Heteroscedasticity

A sequence of random variables is said to be heteroscedastic, if the random variables have different variances. Heteroskedasticity is one of the most important concept in finance. It is connected with the fact that market returns of an individual stock or index returns, returns of commodity and energy markets almost always exhibit heteroskedasticity.

Heteroskedasticity can be one of two following forms:

- A process is said to be *unconditionally heteroscedastic* if unconditional variances are not constant. It is known that stock or bond returns demonstrate heteroscedastic behavior. The prices exhibit non-constant volatility, but periods of low or high volatility are generally not known in advance.

- A process is said to be *conditionally heteroscedastic* if conditional variances are not constant. Oil prices exhibit unconditional heteroscedasticity. The prices tend to have higher volatilities during the Summer than during other seasons.

If a process is unconditionally heteroscedastic, then it is necessarily conditionally heteroscedastic. The converse is not true. If a process is not unconditionally heteroscedastic or not conditionally heteroscedastic, it is said to be unconditionally homoscedastic or conditionally homoscedastic, respectively.

In finance, a variety of models are used for conditionally heteroscedastic processes. These include

- autoregressive conditional heteroscedastic (ARCH) models;
- generalized ARCH (GARCH) models
- regime-switching models; and
- stochastic volatility models.

White test White [1980] is one of well-known methods to test for the presence of heteroscedasticity.

One can use also Engle Test for Conditional Heteroscedasticity (see e.g. Tsay [2005]).

4.3 Risk estimation

4.3.1 Risk measures

Risk measures:

1. Value-at-Risk;
2. Tail conditional expectation;
3. Expected shortfall;
4. Entropic risk measure;
5. Superhedging price.

One of the most well-known risk measures is Value-at-Risk.

Given $\alpha \in (0,1]$, the real number q is said to be α-*quantile* of the random variable X under the probability distribution P if one of the three properties is satisfied:

1. $P(X \leq q) \geq \alpha \geq 1 - P(X \leq q)$;
2. $P(X \leq q) \geq \alpha$ and $P(x \geq q) \geq 1 - \alpha$;
3. $F_X(q) \geq \alpha$ and $F_X(q^-) \leq \alpha$, where F_X is the cumulative distribution function of X, and $F_X(q^-) = \lim_{X \to q,\ X<q} F(x)$.

Jorion (Jorion [2001]) defines Value-at-Risk of an asset as ”the quantile of the projected distribution of gains and losses over the target horizon. If α is selected confidence level, VaR corresponds to the $1 - \alpha$ lower-tail level”:

Definition 1 *Given $\alpha \in (0,1]$, we define the Value-at-Risk (VaR) at level α of the random X with distribution P as negative of the quantile q_α^+ of X, i.e.*

$$VaR_\alpha(X) = -\inf\{x : \ P(X \leq x) \geq \alpha\}.$$

For instance, with 95% confidence level, VaR should be such that it exceeds 5% of the total number of observations in the distribution.

Even though VaR is one of the most commonly used risk measures, it does not satisfy certain properties which are considered to be desirable in any risk measure. These properties are described next.

4.3.2 Properties of coherent risk measures

Denote $\Omega_t = \{\omega\}$ the set of all states of the world at the moment t. We suppose that all elements of the set Ω at the end of period t is known, but the probabilities of states are unknown.

Let X denotes the final worth (at the moment t) of a position for each element $\omega \in \Omega_t$. X is a random variable. Denote G the set of all risks, i.e. the set of all real-valued functions defined on Ω.

It is supposed that any mapping $\rho : G \to \mathbb{R}$ defined on G with range in $\mathbb{R}$ is *a measure of risk*.

A risk measure satisfying the following properties is called *coherent*.

Axiom 1 *(Monotonicity) For all $X, Y \in G$ such that $X \leq Y$, we have*

$$\rho(X) \geq \rho(Y).$$

Axiom 2 *(Positive homogeneity) For all $\alpha \geq 0$ and all $X \in G$, we have*

$$\rho(\alpha X) = \alpha \rho(X).$$

Axiom 3 *(Subadditivity) For all $X, Y \in G$, we have*

$$\rho(X + Y) \leq \rho(X) + \rho(Y).$$

The property 3 is based on the principle of diversification that states that the risk of portfolio always less or equal to risk of it parts.

Axiom 4 *(Translation invariance) For all $X \in G$ and all $\alpha \in \mathbb{R}$, we have*

$$\rho(X + \alpha) = \rho(X) - \alpha.$$

In particular, if $\alpha = \rho(X)$ then $\rho(X + \rho(X)) = 0$ for every $X \in G$. The property 4 states that adding cash amount to the initial position, one decreasing the risk measure on the same value.

VaR does not satisfy the sub-additivity property and therefore is not coherent risk measure.

In the next subsection, we look at the most common method of computing the Value at Risk.

4.3.3 Variance-Covariance Method

This approach is based on the assumption that the underlying market factors have a (multivariate) normal distribution. Then, under some additional assumptions, it is possible to find the distribution of assets (or portfolio) profits and losses. The distribution will be normal as well. Then we can use standard properties of the normal distribution (see (3.3)) to find the loss that equals or exceeds α percent of the time, i.e. $VaR(\alpha)$.

If a probability of α% is used to find 1-day VaR, then VaR is equal to $f(\alpha)$ times the standard deviation of changes in portfolio value:

$$VaR_t(\alpha) = M_{t-1} f(\alpha) \sigma_t,$$

where M_{t-1} is amount of money invested at the moment $t - 1$, $f(\alpha)$ is defined in Table 4.1, and σ_t is estimated volatility (standard deviation) at the moment t.

Table 4.1: The values of the multiplication factor at the formula for VaR estimation for different confidence levels

α	0.5%	1%	2.5%	5%	10%	25%
$f(\alpha)$	2.5758	2.3263	1.9600	1.6449	1.2816	0.6745

4.3.4 Capital Requirements for Market Risk

Despite that VaR models have some shortcomings, they have been accepted by banking regulators as a tools for calculating capital requirements for market risk. The current regulatory framework was set up in the 1988 Basle

Capital Adequacy Accord. The document presents minimum capital requirements for banks' credit risk exposure.

In 1996 American bank regulatory agency release a settlement to the Accord. One of the approach proposed in the settlement (known as "internal models") is based on the Value-at-Risk estimates calculated on the basis of bank's internal risk measurement model using VaR at level 0.01 with 10-day horizon.

According to the settlement, a bank's market risk capital requirement at day t, MRC_t, has to be

$$MRC_t = S_t \max\left\{\frac{1}{60}\sum_{i=1}^{60} VaR_{t-i}(10,1), VaR_t(10,1)\right\} + SR_t,$$

where

- $VaR_i(10,1)$ denotes VaR at level 0.01 with 10-day horizon at the day i,
- S_t is a regulatory multiplication factor, which depends on the accuracy of bank's VaR model. It calculates based on the number of times that daily trading losses exceed the corresponding bank's VaR estimates over the last 250 trading days (see Table 4.2).
- SR_t is an additional capital charge.

Table 4.2: The values of the regulatory multiplication factor

Number of Exceptions (Out of 250 Trading Day	0-4	5	6	7	8	9	10 and more
Scaling Factor, S_t	3.00	3.40	3.50	3.65	3.75	3.85	4

Chapter 5

The Family of ARCH and GARCH models

5.1 Introduction

It is well-known that financial markets and investors react nervously to important news, economic crises, wars, political disorders or natural disasters. In such periods prices of financial assets may fluctuate very much. It means that the conditional variance for the given past

$$Var(X_t|I_{t-1}) := Var(X_t|X_{t-1}, X_{t-2}, \ldots)$$

is not constant over time and the process X_t is conditionally heteroscedastic. In the other words, volatility

$$\sigma_t = \sqrt{Var(X_t|I_{t-1})}$$

changes over time. Understanding the nature of such time dependence is very important for many macroeconomic and financial applications, e.g. irreversible investments, option pricing, asset pricing etc. Models of conditional heteroscedasticity for time series have a very important role in todayís financial risk management and its attempts to make financial decisions on the basis of the observed price asset data P_t in discrete time. Prices P_t are believed to be non-stationary so they are usually transformed in the so-called log–returns

$$X_t = \log P_t - \log P_{t-1}$$

Log returns are supposed to be stationary, at least in periods of time that are not too long. Very often in the past it was suggested that (X_t) represents a sequence of independent identically distributed random variable, in other words, that prices evolve like a random walk. Samuelson suggested modeling speculative prices in the continuous time with the geometric Brownian motion. Discretization of that model leads to a random walk with independent identically distributed Gaussian increments of log return prices in discrete time. This hypothesis was rejected in the early sixties. Empirical studies based on the log return time series data of some US stocks showed the following observations, the so-called stylized facts of financial data:

1. serial dependence are present in the data;
2. volatility changes over time;
3. distribution of the data is heavy-tailed, asymmetric and therefore not Gaussian.

These observations clearly show that a random walk with Gaussian increments is not a very realistic model for financial data. It took some time before R. Engle found a discrete model that described very well the previously mentioned stylized facts of financial data, but it was also relatively simple and stationary so the inference was possible. The ARCH model was introduced by Engle [1982]. Engle called his model autoregressive conditionally heteroscedastic – ARCH, because the conditional variance (squared volatility) is not constant over time and shows autoregressive structure. This model is a convenient way of modeling time-dependent conditional variance. Some years later, Bollerslev [1986] generalized this model as the GARCH model (Generalized Autoregressive Conditional Heteroscedasticity).

5.2 ARCH Model

The primary idea is to use the rolling standard deviation to estimate the variance of the return. This is the standard deviation calculated using a fixed number of the most recent observations, for example 22 business days of data. It assumes that the variance of tomorrow's return can be describe as an equally weighted average of the squared residuals from the last 22 days. The assumption of equal weights seems unrealistic, since the more recent events must be more relevant and therefore should have higher weights. The ARCH model proposed by Engle [1982] let these weights be parameters to be estimated. Thus, the model allowed the data to determine the best weights to use in forecasting the variance.

Let X_t be the log return of a particular stock or the market portfolio from time $t-1$ to time t. Let I_{t-1} denotes the past information set containing the realized values of all relevant variables up to time $t-1$.

Suppose investors know the information in I_{t-1} when they make their investment decision at time $t-1$. Then the relevant expected return μ_t to the investors is the conditional expected value of X_t, given I_{t-1}, i.e.

$$\mu_t = E(X_t|I_{t-1}).$$

The relevant expected volatility σ_t^2 to the investors is conditional variance of X_t, given I_{t-1}, i.e.

$$\sigma_t^2 = Var(X_t|I_{t-1}).$$

Then

$$\epsilon_t = X_t - \mu_t$$

is the unexpected return at time t.

In the paper Engle and Ng [1993] ϵ_t is treated as a collective measure of news at time t. A positive ϵ_t (an unexpected increase in price) suggests the arrival of good news, while a negative ϵ_t (an unexpected decrease in price) suggests the arrival of bad news. Further, a large value of $|\epsilon_t|$ implies that the news is "significant" or "big" in the sense that it produces a large unexpected change in price.

Engle Engle [1982] suggests that the conditional variance σ_t^2 can be modeled as a function of the lagged ϵ's. That is, the predictable volatility is dependent on past news. The most detailed model he develops is the q-th order autoregressive conditional heteroscedasticity model, the ARCH(q).

Let (u_n) be a sequence of i.i.d. random variables such that $u_t \sim N(0,1)$. A process (ϵ_t) is said to be the ARCH(q) process if

$$\epsilon_t = \sigma_t u_t, t \in \mathbb{Z}, \tag{5.1}$$

where (σ_t) is a nonnegative process such that

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2,$$

where $\alpha_0, \alpha_1, \ldots, \alpha_q$ are constant parameters. The effect of a return shock i periods ago ($i \leq q$) on current volatility is governed by the parameter α_i. Normally, one would expect that $\alpha_i < \alpha_j$ for $i > j \geq 1$. That is, the older the news, the less effect it has on current volatility. In an ARCH(q) model, old news which arrived at the market more than q periods ago has no effect at all on current volatility.

5.3 GARCH(p, q) Model

5.3.1 Model Description

Bollerslev (Bollerslev [1986]) generalizes the ARCH(q) model to the GARCH(p, q) model in the following way.

Let (u_n) be a sequence of i.i.d. random variables such that $u_t \sim N(0, 1)$.

A process (ϵ_t) is said to be the generalized autoregressive conditionally heteroscedastic or GARCH(p, q) process if

$$\epsilon_t = \sigma_t u_t, t \in \mathbb{Z}, \tag{5.2}$$

where (σ_t) is a nonnegative process such that

$$\begin{aligned} \sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \ldots + \alpha_q \epsilon_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \ldots + \beta_p \sigma_{t-p}^2 = \\ = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \end{aligned} \tag{5.3}$$

and

$$\alpha_i > 0,\ i = 0, \ldots, q,\ \beta_j > 0,\ j = 1, \ldots, p. \tag{5.4}$$

The conditions (5.4) on parameters ensure strong positivity of the conditional variance (5.3).

Let B denotes the lag-operator, i.e.

$$By_t = y_{t-1},\ B^k = B(B^{k-1}).$$

Then we can write (5.3) in the following way:

$$\sigma_t^2 = \alpha_0 + \alpha(B)\epsilon_t^2 + \beta(B)\sigma_t^2 \tag{5.5}$$

where

$$\alpha(B) = \alpha_1 B + \alpha_2 B^2 + \ldots + \alpha_q B^q,$$
$$\beta(B) = \beta_1 B + \beta_2 B^2 + \ldots + \beta_p B^p.$$

Suppose that the roots of the characteristic equation

$$1 - \beta_1 x - \beta_2 x^2 - \ldots - \beta_p x^p = 0$$

lie outside the unit circle and the process (ϵ_t) is stationary. Then we can write (5.3) as

$$\sigma_t^2 = \frac{\alpha_0}{1-\beta(B)} + \frac{\alpha(B)}{1-\beta(B)}\epsilon_t^2 = \alpha_0^* + \sum_{i=1}^{\infty} \delta_t \epsilon_{t-i}^2, \tag{5.6}$$

where $\alpha_0^* = \frac{\alpha_0}{1-\beta_1}$, and δ_i are coefficient of B^i in expansion of $\frac{\alpha(B)}{1-\beta(B)}$.

It follows from (5.6) that the GARCH(p, q) process can be considered as an ARCH process of infinite order with a fractional structure of the coefficients.

From (5.2) it is obvious that the GARCH(1,1) process is weakly stationary if the process (σ_t^2) is weakly stationary. So if we want to study the properties and higher order moments of GARCH(1,1) process it is enough to do so for the process (σ_t^2).

It was mentioned by Bollerslev [1992] GARCH(1,1) is more preferable in most cases as compared to GARCH(p, q). We recall that a process (ϵ_t) is said to be the generalized autoregressive conditionally heteroscedastic or GARCH(1,1) process if $\epsilon_t = \sigma_t u_t, t \in \mathbb{Z}$, where (σ_t) is a nonnegative process such that

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2.$$

In the model, α reflects the influence of random deviations in the previous period on σ_t, whereas β measures the part of the realized variance in the previous period that is carried over into the current period. The sizes of the parameters α and β determine the short-run dynamics of the resulting volatility time series. Large GARCH error coefficients, α, mean that volatility reacts intensely to market movements. Large GARCH lag coefficients, β, indicate that shocks to conditional variance take a long time die out, so volatility is persistent.

We will use a GARCH model of order 1 since it has been shown to provide a abstemious representation of the conditional variance. GARCH(1,1) was carefully tested (see e.g. the survey Bollerslev [1992]). Also Hansen and Lunde (Hansen and Lunde [2001]) "found no evidence that a GARCH(1,1) is outperformed by more sophisticated models" for prediction of variance of stock returns.

5.3.2 Maximum likelihood estimation of GARCH model

To calibrate the GARCH(p, q) model we can use different methods including the least-square estimator or generalized moment method, but in this work we will apply the maximum likelihood approach.

The subsection describes quasi-maximum likelihood estimation (QML) of model (5.2), (5.3), (5.4). The vector of model parameters is

$$\theta = (\alpha_0, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p)^T.$$

Since the errors are assumed to be conditionally i.i.d. normal, maximum likelihood is a natural choice to estimate the unknown parameters, θ. We will assume that θ belongs to the set

$$\Theta := \{(\alpha_0, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p)^T : \ \alpha_0 \geq 0, \ \alpha_i > 0, \ \beta_j > 0\}.$$

Denote

$$\theta^* = (\alpha_0^*, \alpha_1^*, \ldots, \alpha_q^*, \beta_1^*, \ldots, \beta_p^*)^T$$

the vector of the true values of parameters. The aim is to find θ^* that maximize a QML function given an observation sequence

$$\epsilon_0, \ldots, \epsilon_n$$

of length n.

Define the sequence $(\tilde{\sigma}_1, \ldots, \tilde{\sigma}_n)$ by recursion:

$$\tilde{\sigma}_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \tilde{\sigma}_{t-j}^2,\ 1 \leq t \leq n,$$

where $\epsilon_{1-q}, \ldots, \epsilon_0$ and $\tilde{\sigma}_{1-p}, \ldots, \tilde{\sigma}_0$ are an initial values of ϵ's and σ's respectively.

Given the initial values, the Gaussian quasi-likelihood function can be written as follows

$$L_n(\theta) = L_n(\theta; \epsilon_1, \ldots, \epsilon_n) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi\tilde{\sigma}_t^2}} \exp\left(-\frac{\epsilon_t^2}{2\tilde{\sigma}_t^2}.\right)$$

The optimal estimation of θ is defined by

$$\tilde{\theta} = \arg\max_{\theta\in\Theta} L_n(\theta) = \arg\max_{\theta\in\Theta} F_n(\theta),$$

where

$$F_n(\theta) := -\sum_{t=1}^{n} \left(\frac{\epsilon_t^2}{\tilde{\sigma}_t^2} + \log \tilde{\sigma}_t^2\right)$$

is log quasi-likelihood function (constant terms are ignored).

MATLAB code one can find in Appendix B. Standard error of estimates of parameters were obtained using the statistical environment R.

- Function GARCHrun.m
  1. loads the file with input data;
  2. runs function GARCHcalibration.
- Function GARCHcalibration is a function with two input parameters
  1. array of input data;
  2. initial values of model parameters.

  Function GARCHcalibration uses iteratively the functions garchMaxlikelihood and MATLAB function fminsearch to find the optimal values of model parameters.
- Function garchMaxlikelihood has two input arguments

1. array of input data;
2. values of model parameters omega, alpha, beta.

The output of the function is the value of maximum likelihood function of the GARCH model with given parameters and input data.

5.4 Asymmetric GARCH models

The asymmetric effect was mentioned for the first time by Black in the paper Black [1976] and was also studied by Kenneth R. French, G. William Schwert and Robert F. Stambaugh (French et al. [1987]), D. Nelson (Nelson [1990]), and G. William Schwert (Schwert [1990]).

To capture the asymmetric effects a few models were introduced since then. In this section we will describe the the following most commonly used asymmetric GARCH models:

1. Exponential GARCH (EGARCH) model;
2. The Quadratic GARCH (QGARCH) model;
3. The Glosten-Jagannathan-Runkle GARCH (GJR-GARCH) model;
4. The Threshold GARCH (TGARCH) model.

We will consider the problem of calibration of this models as well.

5.4.1 EGARCH Model

Exponential GARCH (EGARCH) model can be defined by D. Nelson as follows Nelson [1991]:

$$\log(\sigma_t^2) = \alpha_0 + \sum_{i=1}^{q} \left(\gamma_i \frac{\epsilon_{t-i}}{\sigma_{t-i}} + \alpha_i \left| \frac{\epsilon_{t-i}}{\sigma_{t-i}} \right| \right) + \sum_{j=1}^{p} \beta_j \log(\sigma_{t-j}^2),$$

where α_0, α_i, β_j, γ_i are parameters.

The EGARCH model is asymmetric because the levels of $\frac{\epsilon_{t-i}}{\sigma_{t-i}}$'s are included with coefficients γ_i, which are typically negative. Thus positive return shocks generate less volatility than negative return shocks.

The paper Engle and Ng [1993] pointed out that

1. The EGARCH model allows good news and bad news to have a different impact on volatility, while the standard GARCH model does not;
2. The EGARCH model allows big news to have a greater impact on volatility than the standard GARCH model.

EGARCH(1,1) model can be written as

$$\log(\sigma_t^2) = \alpha_0 + \beta \log(\sigma_{t-1}^2) + \gamma \frac{\epsilon_{t-1}}{\sigma_{t-1}} + \alpha_1 \frac{|\epsilon_{t-1}|}{\sigma_{t-1}},$$

where α_0, α_1, β, γ are parameters.

5.4.2 QGARCH model

The Quadratic GARCH (QGARCH) model was invented and studied by Sentana in 1995 Sentana [1995]. The model captures the asymmetric effects of positive and negative shocks. QGARCH(1,1) model may be defined as follows.

Let (u_n) be a sequence of i.i.d. random variables such that $u_t \sim N(0,1)$.

A process (ϵ_t) is said to be QGARCH(1,1) process if

$$\epsilon_t = \sigma_t u_t, t \in \mathbb{Z}, \tag{5.7}$$

where (σ_t) is a nonnegative process such that

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \gamma_1 \epsilon_{t-1} \tag{5.8}$$

and

$$\alpha_0, \alpha_1, \beta_1, \gamma_1 > 0. \tag{5.9}$$

Note that QGARCH model can be reduced to the well-studied GARCH(1,1) model for $\gamma = 0$, and the model captures the leverage effect for $\gamma < 0$.

5.4.3 GJR–GARCH model

The Glosten-Jagannathan-Runkle GARCH (GJR-GARCH) model was developed by Glosten, Jagannathan and Runkle in 1993 Glosten et al. [1993]. The model is similar to QGARCH in the sense of capturing of asymmetry in the ARCH process.

As it is above let (u_n) denotes a sequence of i.i.d. random variables such that $u_t \sim N(0,1)$.

A process (ϵ_t) is said to be GJR–GARCH(1,1) process if

$$\epsilon_t = \sigma_t u_t, t \in \mathbb{Z}, \tag{5.10}$$

where (σ_t) is a nonnegative process such that

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \gamma_1 \epsilon_{t-1}^2 \chi_{t-1}, \tag{5.11}$$

where $\alpha_0, \alpha_1, \beta_1, \gamma_1 > 0$ and

$$\chi_{t-1} = 0 \text{ if } \epsilon_{t-1} \geq 0, \text{ and } \chi_{t-1} = 1 \text{ if } \epsilon_{t-1} < 0. \tag{5.12}$$

5.4.4 Threshold GARCH model

The Threshold GARCH (TGARCH) model was studied by Zakoian in 1994 (Zakoian [1994]). TGARCH model is indeed similar to GJR–GARCH models defined in the previous subsection. The specification of TGARCH model is one on conditional standard deviation instead of conditional variance. Here is definition of the model. A process (ϵ_t) is said to be TGARCH(1,1) process if

$$\epsilon_t = \sigma_t u_t, t \in \mathbb{Z}, \tag{5.13}$$

where (σ_t) is a nonnegative process such that

$$\sigma_t = \alpha_0 + \alpha^+ \epsilon_{t-1}^+ + \alpha^- \epsilon_{t-1}^- + \beta_1 \sigma_{t-1}, \tag{5.14}$$

where

$$\epsilon_{t-1}^+ = \begin{cases} \epsilon_{t-1}, & \text{if } \epsilon_{t-1} \geq 0, \\ 0, & \text{if } \epsilon_{t-1} < 0. \end{cases} \tag{5.15}$$

and

$$\epsilon_{t-1}^- = \begin{cases} \epsilon_{t-1}, & \text{if } \epsilon_{t-1} \leq 0, \\ 0, & \text{if } \epsilon_{t-1} > 0. \end{cases} \tag{5.16}$$

5.4.5 Calibration of Asymmetric Models

Given an observation sequence

$$\epsilon_0, \dots, \epsilon_n$$

of length n, define the sequence $(\tilde{\sigma}_1, \dots, \tilde{\sigma}_n)$

1. in EGARCH(1,1) model by

$$\log(\tilde{\sigma}_t^2) = \alpha_0 + \beta \log(\tilde{\sigma}_{t-1}^2) + \gamma \frac{\epsilon_{t-1}}{\tilde{\sigma}_{t-1}} + \alpha_1 \frac{|\epsilon_{t-1}|}{\tilde{\sigma}_{t-1}},$$

   where α_0, α_1, β, γ are estimated (unknown) parameters;

2. in QGARCH(1,1) model by

$$\tilde{\sigma}_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \tilde{\sigma}_{t-1}^2 + \gamma_1 \epsilon_{t-1}$$

   where $\alpha_0, \alpha_1, \beta_1, \gamma_1 > 0$ are estimated (unknown) parameters;

3. in GJR-GARCH(1,1) model by

$$\tilde{\sigma}_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \tilde{\sigma}_{t-1}^2 + \gamma_1 \epsilon_{t-1}^2 \chi_{t-1},$$

   where $\alpha_0, \alpha_1, \beta_1, \gamma_1 > 0$ are estimated (unknown) parameters and χ_{t-1} is defined in (5.12).

4. in Threshold GARCH(1,1) model by

$$\tilde{\sigma}_t = \alpha_0 + \alpha^+ \epsilon_{t-1}^+ + \alpha^- \epsilon_{t-1}^- + \beta_1 \tilde{\sigma}_{t-1},$$

   where α_0, α^+, α^-, β_1 are estimated (unknown) parameters and ϵ_{t-1}^+, ϵ_{t-1}^- are defined in (5.15), (5.16).

If we assume that the likelihood function is Gaussian, then the log-likelihood function can be written as

$$F_n(\theta) := -\sum_{t=1}^{n} \left(\frac{\epsilon_t^2}{\tilde{\sigma}_t^2} + \log \tilde{\sigma}_t^2 \right)$$

(constant terms are ignored). The maximum likelihood estimator of θ is defined by

$$\tilde{\theta} = \arg\max_{\theta \in \Theta} F_n(\theta).$$

It is worth noting that estimates obtained by maximizing the log likelihood of a normal distribution are strongly consistent (although they are not efficient). Recall that an estimator has the property of strong consistency if parameter estimates converge to the true parameters (even assuming the wrong conditional distribution).

If the GARCH model correspond the true data process, then the parameters of the GARCH model are chosen such that the conditional expectation of the generalized error 0. The normal distribution has the property that these parameters will correspond to those of the original data process even if the conditional distribution is incorrect.

The assumption that the errors are conditionally normal has some advantages: estimation is quite simple and parameters are consistent for the true parameters. But the alternative (non–normal) distributions are more useful in application to Value-at-Risk in which case the choice of density may lead to a better prediction capacity.

Some researchers estimate GARCH models assuming an alternative distribution (see, for example Bollerslev [1987]). It gave a better approximation to the conditional distribution of the standardized returns. Moreover, in the case of MLE, the estimates are fully efficient.

5.5 Empirical Study

5.5.1 Data Description

The data set we analyzed in this work is stock market closing daily prices of the General Electric Company (GEC.L)

The sample period is from January 2, 2008 to December 31, 2010. Total number of observations is 757. Data set are taken from UK Stock Market FTSE100 and downloaded from Yahoo!Finance site. The sample is divided in twelve parts for two purposes:

- *In-sample estimation procedure.* The twelve parts of observations are used as in-sample data. The first data set consists of data from January

2, 2008 to December 31, 2009. The i-th data set represents data from the first day of i-th month of 2008 to the last day of $i-1$-th month of 2010. The twelfth data set is from January 2, 2009 to December 31, 2010. We will estimate parameters of models on each of this data sets. In-sample data sets are listed in Table A.

- *Out-of-sample evaluation*. The remaining observations (from January 2, 2010 to December 31, 2010) are used as out-of-sample for forecast evaluation purposes. We will use the parameters obtained on previous step to estimate predictive properties of models. We divide the data set (from January 2, 2010 to December 31, 2010) on twelve one-month length parts. The data of i-th month of 2010 will be used to estimate the forecast properties of model which parameters was obtained based on i-th in-sample data set. Out-of-sample data sets are determined in Table A.2.

Table 5.1 shows some descriptive statistics of stock prices of the GE company. Figure 5.1 presents the General Electric prices movement during the period from January 2, 2008 to December 31, 2010.

Table 5.1: Descriptive statistics of General Electric company stock market closing daily prices (January 2, 2008 - December 31, 2010)

Form	To	Size	Mean	Std. dev.	Min	Max
02/01/2008	31/12/2010	757	19.29868	7.514138	6.66	38.43

Due to significant changes in the level of prices from date to date, it is more appropriate to base volatility measures on percentage return, rather than absolute price movement. Returns are calculated by

$$r_t = 100 \log \frac{p_t}{p_{t-1}},$$

where p_t and P_{t-1} are closing price on dates t and $t-1$ respectively. The reason for using natural logarithm of trading volume is to improve the normality of the series in order to better fit into the GARCH-type models.

The descriptive statistics of GE stock log returns in Table A.3 shows the stylized fact that the returns series is not close to normality as reflected by kurtosis value. Returns ranges from -13.68411 to 17.98444 and mean of returns is not statistically different from zero at 5% level. Historical movement of

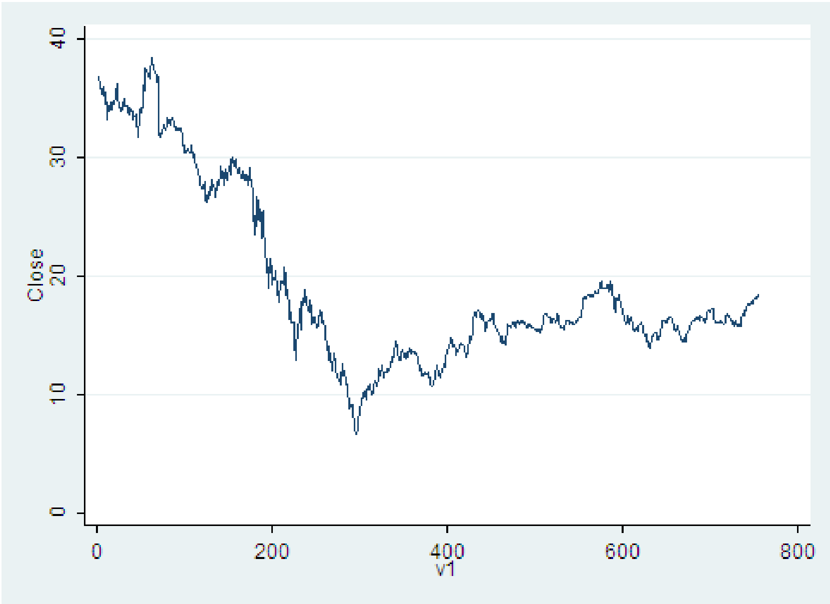

Figure 5.1: Historical movement of General Electric stock market closing daily prices (January 2, 2008 - December 31, 2010)

the GE log returns one can find in Figure 5.2. Time series of returns appears to show the signs of ARCH effects in that the amplitude of the returns varies over time. Figure 5.4 presents the histogram of the GE log returns during the period.

The positive skewness implies that returns distribution has a higher probability of earning positive returns. Kurtosis value is 7.708553 and this is much larger than 3. This shows that for our returns series, the distribution has fatter tails and sharper peaks at the center compared to the normal distribution (see Fig. 5.4). That behavior is known to occur in financial markets often.

The Jarque–Bera test statistics was also found to have a high value at significant 1% level and therefore we have to reject the null hypothesis of normality. Shapiro–Wilk W test also rejects the hypothesis of normal distribution of the data.

We used the test of the white noise process given by the Ljun-Box-Pierce portmanteau test statistics in order to test the hypothesis of independence. The first-order autocorrelations of our series are small and statistically non-

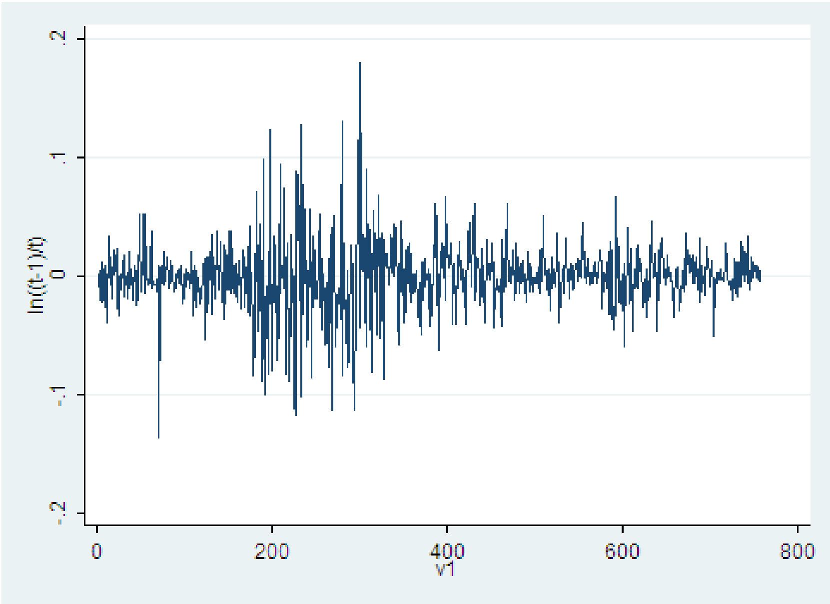

Figure 5.2: Historical movement of log returns of the General Electric stock market closing daily prices (January 2, 2008 - December 31, 2010)

significant and returns do not have first-order autocorrelations. The computed statistical values of Ljung-Box-Pierce for lags 10 and 20 for our returns series (denoted by Q(10) and Q(20) respectively) are relatively large. The Q(10) and Q(20) test statistic reject the null hypothesis of uncorrelated price changes, suggesting a slowly decaying autoregressive effect. Thus, the null hypothesis of strict white noise is rejected.

5.5.2 Empirical results

Kupiec's test.

Since the late 1990's a variety of tests have been proposed that can be used to estimate the accuracy of a VaR model. To compare the performance of models, in this work we will use the approach developed by Kupiec Kupiec [1995].

The idea is quite simple. This tests are concerned with whether or not the reported VaR is violated more (or less) than $\alpha 100\%$ of the time. Kupiec

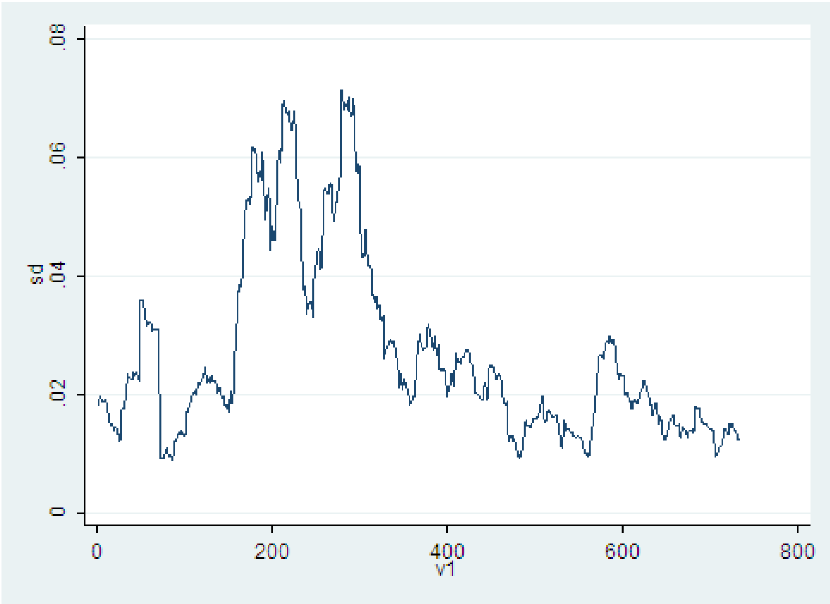

Figure 5.3: The rolling standard deviation of log returns of the General Electric stock market closing daily prices (January 2, 2008 - December 31, 2010), with 21 days rolling window

finds a proportion of failures that examines how many times an asset's VaR is violated over a given span of time. If the number of violations differs considerably from $\alpha 100\%$ of the sample, then the accuracy of the underlying risk model is called into question.

Denote by y_t the actual losses at the day t. Consider the event that the loss on a portfolio exceeds its reported VaR (estimated by a model), VaR_t, i.e. $y_t > VaR_t$. Let $\{I_t\}$, $1 \leq t \leq T$, be the sequence with

$$I_t = \begin{cases} 1, & \text{if } y_t > VaR_t, \\ 0, & \text{if } y_t \leq VaR_t. \end{cases} \tag{5.17}$$

Denote by

$$N = \sum_{t=1}^{T} I_t \tag{5.18}$$

the number of such violations.

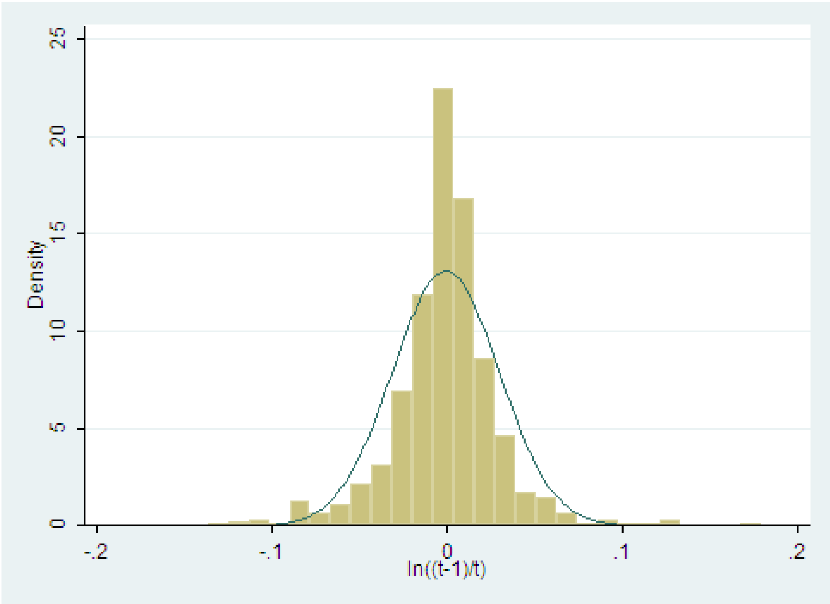

Figure 5.4: Histogram of log returns of the General Electric stock market closing daily prices (January 2, 2008 - December 31, 2010)

Kupiec's test statistic has the form Kupiec [1995]

$$L = 2log\left(\left(\frac{1 - N/T}{1-\alpha}\right)^{T-N}\left(\frac{N/T}{\alpha}\right)^{N}\right)$$

The intervals of non-rejection for some α is given in Table 5.2.

Table 5.2: The intervals of non-rejection regions for Kupiec's test

α	$T = 250$	$T = 500$	$T = 750$	$T = 1000$
5%	$7 \leq N \leq 19$	$17 \leq N \leq 35$	$25 \leq N \leq 49$	$38 \leq N \leq 64$
1%	$1 \leq N \leq 6$	$2 \leq N \leq 9$	$3 \leq N \leq 13$	$5 \leq N \leq 16$

Christoffersen Christoffersen [1998] points out that the problem of determining the accuracy of a VaR model can be reduced to the problem of determining whether the sequence of violations, $\{I_t\}$, $1 \leq t \leq T$, satisfies two properties:

1. *Unconditional coverage property*. The unconditional coverage property

places a restriction on how often VaR violations may occur. The probability of realizing a loss in excess of the reported VaR_t must be precisely $\alpha100\%$ or in terms of the previous notation, $P(I_{t+1}(\alpha)) = \alpha$. If it is the case that losses in excess of the reported VaR occur more frequently than 100% of the time then this would suggest that the reported VaR measure systematically understates the portfolio's actual level of risk. The opposite finding of too few VaR violations would alternatively signal an overly conservative VaR measure.

2. *Independence property*. The independence property places a strong restriction on the ways in which these violations may occur. Specifically, any two elements of the hit sequence, $I_{t+j}(\alpha), I_{t+k}(\alpha)$ must be independent from each other.

The group of competing GARCH models (ARCH, GARCH, TGARCH, GJR-GARCH) are estimated using quasi-maximum likelihood method.

Table A.4 presents autocorrelations, partial autocorrelations, and Portmanteau (Q) statistics. It shows that there is no autocorrelation in GE stock returns time series. Therefore, we do not include autoregressive and moving average terms in mean equation. We will assume

$$\mu = \mathbb{E}(r_t).$$

The MATLAB code one can find in Appendix B.

Standard error of estimates of parameters were obtained using the statistical environment R.

In Appendix A one can find

- ARCH(1) model parameters's estimation of returns for the GE stock market closing daily prices (January 2, 2008 - December 31, 2010), $\sigma_t^2 = \alpha_0 + \alpha_1\epsilon_{t-1}^2$, Table A.5 and Table A.6;
- GARCH(1,1) model parameters's estimation of returns of the GE company closing daily prices (January 2, 2008 - December 31, 2010, $\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$, Table A.7 and Table A.8;

- Threshold GARCH(1,1) model parameters's estimation of returns of the GE company closing daily prices (January 2, 2008 - December 31, 2010), $\sigma_t = \alpha_0 + \alpha^+\epsilon^+_{t-1} + \alpha^-\epsilon^-_{t-1} + \beta_1\sigma_{t-1}$, Table A.9;
- GJR–GARCH(1,1) model parameters's estimation of returns of the GE company closing daily prices (January 2, 2008 - December 31, 2010), $\sigma^2_t = \alpha_0 + \alpha_1\epsilon^2_{t-1} + \beta_1\sigma^2_{t-1} + \gamma_1\epsilon^2_{t-1}\chi_{t-1}$, where $\chi_{t-1} = 0$ if $\epsilon_{t-1} \geq 0$, and $\chi_{t-1} = 1$ if $\epsilon_{t-1} < 0$, Table A.10 and Table A.11.

Notice that the coefficients sum up to a number less than one, which is required to have a mean reverting variance process. Since the sum is very close to one, this process only mean reverts slowly.

Figure shows the values of conditional variance of General Electric stock market closing daily prices, based on 1-year GARCH model, 1-day prediction.

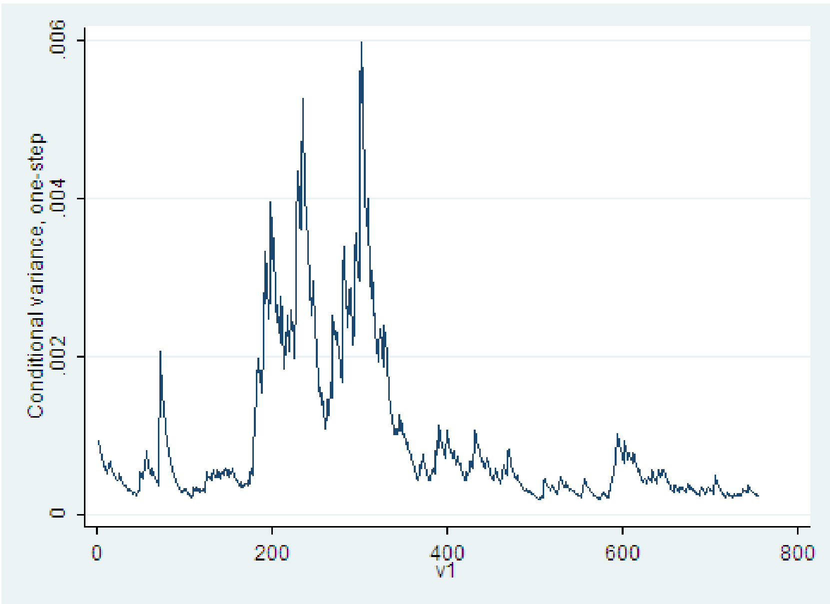

Figure 5.5: Prediction of conditional variance of General Electric stock market closing daily prices (January 2, 2008 - December 31, 2009), based on 1-year GARCH model

Thus, Tables A.5, A.6, A.7, A.8, A.9, A.10, A.11 present estimated parameters for ARCH, GARCH, TGARCH, GJR-GARCH models for all 24 in-sample datasets (Table A). Using these estimated parameters, we can forecast the one day ahead volatility for each day of out-of-sample datasets (Table A.2). To do so we put the values of the estimated parameters in the formula for their respective models. For example, to predict the value of σ_t using Threshold GARCH(1,1) model we have to substitute the estimated values of parameters α_0, α^+, α^- and β_1 in the equation of $TGARCH(1,1)$ model at time $t = k$, $1 \leq k \leq 21$:

$$\sigma_k = \alpha_0 + \alpha^+ \epsilon_{k-1}^+ + \alpha^- \epsilon_{k-1}^- + \beta_1 \sigma_{k-1},$$

where ϵ_{k-1}^+ and ϵ_{k-1}^- are positive and negative part of the value of the squared return of the previous day respectively. Further the value of variance σ_k at time k is used to forecast the volatility at time $k+1$ using out-of-sample data set (out-of-sample data sets are determined in Table A.2):

$$\sigma_{k+1} = \alpha_0 + \alpha^+ \epsilon_k^+ + \alpha^- \epsilon_k^- + \beta_1 \sigma_k.$$

Thus, we can forecast volatility for up to 21st day. The same way is used for the volatility forecast by $ARCH(1,1)$, $GARCH(1,1)$, $GJR - GARCH(1,1)$ models.

The same procedure we use for volatility forecast for each of 12 out-of-sample datasets. Hence volatility is forecasted for 251 days from 04/01/10 to 31/12/10 using each of 4 models, for all of 12 out-of-sample datasets.

Then $VaR_t(\alpha)$, $506 \leq t \leq 757$ is calculated by

$$VaR_t(\alpha) = f(\alpha) \cdot \sigma_t,$$

where $\alpha = 0.05$ or 0.01 is confidence level and σ_t is estimated volatility for the day t.

Thus VaR is forecasted using different models for the whole of the out of sample period of 251 days from date 04/01/10 to 31/12/10.

The values of $VaR_t(\alpha)$ can be used for the purpose of judging the performance of the volatility models. To apply the unconditional coverage of Kupiec Kupiec [1995] it is necessary to find the number of violations by using (5.17). Non-rejection region one can find in the Table 5.2.

Number of VaR violations of each model for 1-year length of in-sample datasets is given in Table 5.3 and Table 5.4. Number of VaR violations of each model for 2-year length of in-sample datasets is given in Table 5.5 and Table 5.6.

Table 5.3 and Table 5.4 show that for 95% and 99% confidence level respectively, all models calibrated on 1-year dataset cannot be rejected.

Table 5.3: Numbers of failures with 95% confidence level, 1-year length of dataset

Model	Length	Conf	Failures	Freq	%
ARCH(1,1)	1 year	95%	8	251	3.17
GARCH(1,1)	1 year	95%	9	251	3.57
GJR-GARCH	1 year	95%	10	251	3.98
TGARCH	1 year	95%	10	251	3.98

Table 5.4: Numbers of failures with 99% confidence level, 1-year length of dataset

Model	Length	Conf	Failures	Freq	%
ARCH(1,1)	1 year	99%	2	251	0.79
GARCH(1,1)	1 year	99%	3	251	1.19
GJR-GARCH	1 year	99%	3	251	1.19
TGARCH	1 year	99%	3	251	1.19

Table 5.5: Numbers of failures with 95% confidence level, 2-year length of in-sample dataset

Model	Length	Conf	Failures	Freq	%
ARCH(1,1)	2 year	95%	1	251	0.40
GARCH(1,1)	2year	95%	1	251	0.40
GJR-GARCH	2 year	95%	1	251	0.40
TGARCH	2 year	95%	2	251	0.79

On the other hand, Table 5.5 and Table 5.6 show that for 95% and 99% confidence level, all models calibrated on 2-year dataset must be rejected, since they over predict the risk and the models falls out of the non-rejection region.

Christoffersen' s interval forecast test

Table 5.6: Numbers of failures with 99% confidence level, 2-year length of in-sample dataset

Model	Length	Conf	Failures	Freq	%
ARCH(1,1)	2 year	99%	0	251	0.00
GARCH(1,1)	2year	99%	0	251	0.00
GJR-GARCH	2 year	99%	0	251	0.00
TGARCH	2 year	99%	0	251	0.00

Christoffersen (1998) proposed a test of conditional coverage. He extends the Kupiec's test to examines whether the probability of an exception on any day depends on the outcome of the previous day.

Thus Christoffersen's test not only covers the violation rate but the independence of exception also. If the model is accurate, then an exception today should not depend on whether or not an exception occurred on the previous day.

Define an indicator variable

$$I_t = \begin{cases} 1, & \text{if violation occurs in the day } t, \\ 0, & \text{if no violation occurs in the day } t, \end{cases}$$

i.e. I_t gets a value of 1 if VaR is exceeded and value of 0 if VaR is not exceeded.

Denote by n_{ij} the number of days when condition $I_t = j$ occurred assuming that condition $I_t = i$ occurred on the previous day.

To illustrate, the outcome can be displayed in a 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

Let

$$\Pi_0 = \frac{n_{01}}{n_{00} + n_{01}}, \ \Pi_1 = \frac{n_{11}}{n_{10} + n_{11}}, \ \Pi = \frac{n_{01} + n_{11}}{n_{00} + n_{01} + n_{10} + n_{11}}$$

Thus, n_{11} presents the number of consecutive exceptions, Π_i is the probability of an exception assuming a state i on the previous day, and Π is the probability of an exception regardless of the previous dayís state. The probabilities are calculated from the observed data.

The test statistic for independence of exception is likelihood-ratio:

$$LR = -2\log\frac{(1-\Pi)^{n_{00}+n_{10}}\Pi^{n_{01}+n_{11}}}{(1-\Pi_0)^{n_{00}}\Pi_0^{n_{01}}(1-\Pi_1)^{n_{10}}\Pi_1^{n_{11}}} \tag{5.19}$$

LR is also $\chi^2(1)$-distributed. If the value of the LR statistic is lower than the critical value of $\chi^2(1)$ distribution, the model passes the test. Higher values lead to rejection of the model.

As an example, consider again the ARCH model results at 95% confidence level. The contingency table can be presented as follows

	$I_{t-1}=0$	$I_{t-1}=1$	
$I_t=0$	235	7	235
$I_t=1$	7	1	8
	235	8	250

In addition, we need to find the probabilities

$$\Pi_0 = \frac{4}{235+7} = 0.032089,\ \Pi_1 = \frac{1}{7+1} = 0.1250,\ \Pi = \frac{7+1}{235+7+7+1} = 0.0320$$

Plugging this data into the likelihood ratio statistic we obtain the test value:

$$LR = -2\log\frac{(1-0.0320)^{235+7}0.0320^{7+1}}{(1-0.032089)^{235}0.032089^{7}(1-0.1250)^{7}0.1250^{1}} = 1.3873$$

The critical value is the 95% percentile of the $\chi^2(1)$ distribution with one degree of freedom, 3.84. As the test statistic value remains below the critical value, the model is accepted.

Table A.12 shows the input data for calculating the LR statistics for each models and confidence level.

Table 5.7 shows that for 95% and 99% confidence level, all models calibrated both on 1-year and on 2-year dataset cannot be rejected.

Table 5.7: Christoffersen's independence test resuls

Model	Length	Conf	LR	Critical Value	
ARCH(1,1)	1 year	95%	1.3873	3.84	Accepted
GARCH(1,1)	1 year	95%	0.6724	3.84	Accepted
GJR-GARCH	1 year	95%	0.8336	3.84	Accepted
TGARCH	1 year	99%	0.8336	6.63	Accepted
ARCH(1,1)	1 year	99%	0.0323	6.63	Accepted
GARCH(1,1)	1 year	99%	0.0729	6.63	Accepted
GJR-GARCH	1 year	99%	0.0729	6.63	Accepted
TGARCH	1 year	99%	0.0729	6.63	Accepted
ARCH(1,1)	2 year	95%	0.0080	3.84	Accepted
GARCH(1,1)	2 year	95%	0.0080	3.84	Accepted
GJR-GARCH	2 year	95%	0.0080	3.84	Accepted
TGARCH	2 year	95%	0.0323	3.84	Accepted
ARCH(1,1)	2 year	99%	0.0000	6.63	Accepted
GARCH(1,1)	2 year	99%	0.0000	6.63	Accepted
GJR-GARCH	2 year	99%	0.0000	6.63	Accepted
TGARCH	2 year	99%	0.0000	6.63	Accepted

5.5.3 Summary

Backtesting procedure shows that the ARCH(1), GARCH (1, 1), TGARCH(1,1) and GJR-GARCH (1, 1) models calibrated on datasets of **1-year length** (January 2, 2009 - December 31, 2009) under the normal distribution performed well forecasting VaR both at 95% and 99% in 2010 year.

However, for 95% and 99% VaR estimations all the ARCH(1), GARCH (1, 1), TGARCH(1,1) and GJR-GARCH (1, 1) models calibrated on datasets of **2-year length** (January 2, 2008 - December 31, 2009) underestimated the risk and should be rejected.

It can be explained by huge difference in the level of volatility in 2008 crisis year compare with the one in 2009 year. The conclusion is that the ARCH(1), GARCH (1, 1), TGARCH(1,1) and GJR-GARCH (1, 1) models are highly sensitive to dataset they are calibrated on.

Chapter 6

Augmented GARCH Models

6.1 News Analytics

6.1.1 Introduction

Many investment companies in the U.S. and Europe have been using news analytics to improve the quality of its business. Interest in news analytics is related to the ability to predict changes of prices, volatility and trading volume on the stock market Tetlock [2007].

News analytics uses some methods and technics of data mining Kantardzic [2003] and relies on methods of computer science, artificial intelligence (including algorithms for natural language processing), financial engineering, mathematical statistics and mathematical modeling. News analytics software signalize traders about the most important events or send their output data directly to automated trading algorithms, which take into account this signals automatically during the trade.

This chapter is a short review of the tools, methods and providers of news analytics and mostly based on the book Mitra and Mitra [2011].

6.1.2 What is the News Analytics?

News analytics can be described as a measurement of the following quantitative and qualitative characteristics of news:

1. **The nature of news** (it determines the impact of news (positive or negative), i.e. how news affects stock prices change; it is believed that positive news about the company leads to a growth in the stock prices of its shares, and negative, on the contrary, can leads to decreasing);
2. **The impact of news** (it is characterized by the influence of news on the scale of the changes caused by the news);
3. **The relevance** (describes how the events, described in a news report, are connected with the trader's interest security);
4. **The novelty** (shows how much news is informative, usually it is inversely correlated with the number of references to events that are written in this news report, with other news).

News analysis is a relatively new tool designed to improve the trading strategies of investors. It is closely connected with the theory of behavioral finance and in some sense, is contrary to the classical economic theory.

Indeed, the famous "efficient markets hypothesis" states Samuelson [1965] that any available information is already reflected in share prices. This condition makes it impossible to attempt to outperform the market in a long period of time through the use of information available on the market. On the other hand, in the modern world, the intensity level of various news agencies is so high (for example, Thomson Reuters has more than 4000 messages per day) that the trader is unable on its own to handle this information flow. Events that are potentially change the situation on the stock exchange, may be lost or omitted in a huge stream of news. In this context, it is unlikely that at any one time all traders will be equally informed of all events affecting the price of certain stocks. That is why the news analytics is an effective tool to gain advantage over other market participants.

Knowing the characteristics of news in numerical indices one can use them in mathematical and statistical models and automated trading systems.

Currently, the tools of the news analytics have been increasingly used by traders in the U.S. and Europe.

The process of news analysis in information systems is automated and usually includes the following steps:

1. collecting news from different sources;
2. preliminary analysis of news;
3. analysis of news-related expectations (sentiments), taking into account the current market situation;
4. designing and using of quantitative models.

The process of news analytics is described in more details in the following sections.

It worth be noted that managers of investment funds rarely use tools of news analytics, since they usually create investment portfolios for a long period of time, and in this case portfolio management does not suggest a frequent resale of securities.

6.1.3 Data Sources

News data can be obtained from various sources:

- **News sources of news agencies.** Until recently, the news had been spread by printed sources, radio, television and it was quite difficult to obtain an overall picture of the news flow. The Internet has changed the process of news analysis; the using of tagging and indexing has made possible their automatic processing.
- **Pre-news** is a raw information material which is used in the preparation of news by reporters. It can be obtained from different primary sources, for example, SEC reports, court documents, reports of various government agencies, business resources, company reports, announcements, industrial, and macroeconomic statistics.

- **Social media** (blogs, social networks, etc.). The quality of news from this type of sources can be vary highly, and this information is often useless. However, you can keep track (evaluate) the mood of a large number of these messages and apply results in trade strategies.

In addition, the financial news can be classified in terms of their expectations. Expected news come out at a scheduled time and often their contents can be predicted on the basis of pre-news. They have a structured format and generally include numeric data, which is convenient for automated analysis (e.g., usually all companies publish annual or quarterly financial reports in the same time). Macroeconomic reports have a strong influence on liquid markets (foreign exchange, futures, government bonds) and are widely used in the automatic trading. Speed and accuracy of processing of such information are important technological requirements. Reports of incomes and losses affect directly the change in stock prices and are widely used in trading strategies.

The main difficulties of the processing of financial information are associated with unexpected news, since the time of their appearance is unknown and, often they have a unstructured text format and do not contain numeric data. They are difficult to process quickly and efficiently, but they may contain information about the causes and consequences of the event. To analyze unexpected news one can use the artificial intelligence systems based on methods of natural language processing.

6.1.4 Pre–Analysis of News

News analytics evaluates the relevance, nature, novelty and the importance of news. The results of processing of news information are used to create signals for investors and traders. These signals can be combined with forecasts from other primary or processed sources.

6.1.5 Providers of News Analytics

In the world there are more than 50 providers of economic news. Bloomberg, Dow Jones and Thomson Reuters are the three largest of them.

About 200 agencies are involved in providing of financial analytics.

The most well-known providers of news analytics and data are:

- **RavenPack** (http://www.ravenpack.com/) is one of the leading providers of real-time news analysis services. The company specializes in linguistic analysis of large volumes of news in real time from news providers. RavenPack News Scores measures the news sentiment and news flow of the global equity market based on all major investable equity securities. News scores include analytics on more than 27,000 companies in 83 countries and covers over 98% of the investable global market. All relevant news items about companies 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 various applications. RavenPack is working with news feeds from the company Dow Jones.

- **Media Sentiment** (www.mediasentiment.com/) has a resource library of nearly 2,000,000 articles and it regularly searches and analyzes output from 6,000+ sources in near-real time to bring investors updated news media sentiment about publicly traded companies, both quickly and effortlessly.

- **Thomson Reuters News Analytics** (http://thomsonreuters.com) automatically analyzes news providing improved buy/hold/sell signals within milliseconds. The system can scan and analyze stories on thousands of companies in real-time and feed the results into your quantitative strategies. With its ability to track news sentiment over time, Thomson Reuters News Analytics provides a more comprehensive understanding of a companyís news coverage, helping to guide trading and investment decisions. It delivers unparalleled insight into a companyís market reputation, giving money managers a unique advantage. *Reuters NewsScope* and *Sentiment Analysis* are new software products, which provide financial news (interest rates, consumer price indices, etc.). These programs are designed for use in automated trading.

In the work we will use the Raven Pack sources of news analytics data. One can found the example of extract from the data in Table 6.1.

Table 6.1: Extracts from Raven Pack news analytics data

TIMESTAMP_UTC:	2010-01-21 21:20:08.297
COMPANY:	JP/7203 (Toyota Motor Corp.)
RP_COMPANY_ID:	CEC128
RELEVANCE:	100
EVENT CATEGORY:	product-recall
EVENT SENTIMENT (ESS):	29
NOVELTY (ENS):	100
NOVELTY ID (ENS_KEY):	D9592AD7D8E718...
COMPOSITE SENTIMENT (CSS):	50
WORD/PHRASE LEVEL (WLE):	50
PROJECTIONS BY COMPANY (PCM):	50
EDITORIALS & COMMENTARY (ECM):	50
REPORTS CORP ACTIONS (RCM):	50
VENTURE, CORPORATE, M& A (VCM):	50
NEWS IMPACT PROJECTION (NIP):	34
RP_STORY_ID:	D9592AD7D8E71...

6.2 Augmented GARCH Models

6.2.1 Description of the Models

Recent studies on the volatility of stock returns have been dominated by time series models of conditional heteroscedasticity and have found strong support for ARCH-GARCH-type effects. However, ARCH-GARCH-type models do not provide a theoretical explanation of volatility or what, if any, the exact contributions of information flows are in the volatility-generating process. One theoretical explanation is the mixture of distribution hypothesis (MDH) advanced by Clark [1973], Epps and Epps [1976], Tauchen and Pitts [1983], and Lamoureax and Lastrapes [1991]. The MDH argues that the variance of returns at a given interval is proportional to the rate of information arrival. As a result, volatility clustering could be a reflection of the serial correlation of information arrival frequencies. All traders simultaneously receive the new price signals and the shift to a new equilibrium is immediate and there will be no intermediate partial equilibrium.

The MDH relies on the following assumptions:

- returns and corresponding trade volumes are jointly independently distributed with finite variance (Harris [1987]);
- the number of events occurring each day is stochastic.

Some of empirical studies, including a pioneer work for the US stock market by Lamoureax and Lastrapes [1990], have found evidence that the inclusion of trading volumes in GARCH models for returns results in a decrease of the estimated persistence, or even causes it to disappear. However, the results from other research indicated that trading volume contributes some information to the returns process, while their results also show persistence in volatility even after incorporate volume effects. These research paper include the futures market by Najand and Yung in Najand and Yung [1991].

If the parameters α, β of GARCH(1,1) model are positive, then shocks to volatility persist over time. The sum $\alpha + \beta$ of these parameters reflects the degree of persistence. Denote by δ_{it} the ith intraday equilibrium price increment in day t. Lamoureux and Lastrapes Lamoureax and Lastrapes [1990] suggest that the innovation upon stock returns is a linear combination of intraday price movements, i.e.,

$$\epsilon_t = \sum_{i=1}^{n_t} \delta_{it},$$

n_t is the number of information flows within day t. Thus, Lamoureux and Lastrapes (Lamoureax and Lastrapes [1990]) consider $\epsilon_t = r_t - \mu_{t-1}$ as an aggregation of price innovations from information flows into the market. Note that they do not differentiate on the type of information flows into the market.

Assume δ_{it} is independent identically distributed with mean zero and variance σ^2 and suppose that n_t is large. Then it follows from the central limit theorem that $\epsilon_t | n_t$ is asymptotically distributed as $\mathcal{N}(0, \sigma^2 n_t)$. Then

$$\Omega_t := \mathbb{E}(\epsilon_t^2 | n_t) = \sigma^2 n_t. \tag{6.1}$$

Suppose that n_t's are serially correlated:

$$n_t = a + b(L) n_{t-1} + u_t, \tag{6.2}$$

where a is a constant, $b(L)$ is a lag polynomial of order q, i.e. $b(L)n_{t-1} = \sum_{k=1}^{q} b_k n_{t-k}$, u_t is a white noise.

Substituting representation (6.2) into equation (6.1), we get

$$\Omega_t = \sigma^2 a + b(L)\Omega_{t-1} + \sigma^2 u_t. \tag{6.3}$$

Equation (6.3) describes the conditional variance of returns is depends on lagged conditional variances and a white noise term. It motivated Lamoureax and Lastrapes [1990] to use additive GARCH-Volume model. As it was mentioned by Lamoureux and Lastrapes Lamoureax and Lastrapes [1990], equation (6.3) captures GARCH-type persistence in conditional variance. However, in the paper (Bauer and Nieuwland [1995]) it was pointed out that the link between (6.3) and GARCH model is not at all straightforward.

The question is what is the best proxy for n_t?

Trading volumes can be good proxies for news arrivals. It follows from the intuitive fact that the more company news arrives, the more investors will interpret the impact of the news differently, and thus the more investors will have an impetus to trade as their expectations about future returns differ.

Lamoureux and Lastrapes (Lamoureax and Lastrapes [1990]) suppose that volume can be considered as a proportional proxy for information arrivals to the market. Volume acts as a mixing variable, i.e. ϵ_t is assumed to be random draws upon alternative distributions, with variances depending upon information available at the time. It leads to a model (6.4), (6.5), where V_t is the volume of trade that occurs in time t.

Following the studies by previous authors, we are trying to estimate four alternative augmented GARCH models of volatility.

We will examine the following alternative GARCH models:

1. *GARCH model augmented with volume.*

   We will consider a process (ϵ_t) such that

$$\epsilon_t = \sigma_t u_t, t \in \mathbb{Z}, \tag{6.4}$$

   where (σ_t) is a nonnegative process such that

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma V_t, \tag{6.5}$$

where $V_t = v_t/v^*$ is the scaled trade volume of the stock at the day t (v_t is daily trade volume of the stock at the day t and $v^* = \max_t\ v_t$), $\omega > 0$, $\alpha, \beta \geq 0$, $\alpha + \beta < 1$ and γ are parameters of the model.

2. *GARCH model augmented with lagged volume.*

   We will consider a process (ϵ_t) such that

$$\epsilon_t = \sigma_t u_t, t \in \mathbb{Z}, \tag{6.6}$$

   where (σ_t) is a nonnegative process such that

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma V_{t-1}, \tag{6.7}$$

   where V_{t-1} is the scaled trade volume for the company in day $t-1$, $\omega > 0$, $\alpha, \beta \geq 0$, $\alpha + \beta < 1$ and γ are parameters of the model.

3. *GARCH model augmented with news intensity.*

   Let r_t and r_t^* denote the log return of the company and log return of FTSE100 index at interval t respectively. We will consider a process $(\epsilon_t) = r_t - (\theta_1 + \theta_2 r_t^*)$ such that

$$\epsilon_t = \sigma_t u_t, t \in \mathbb{Z}, \tag{6.8}$$

   where (σ_t) is a nonnegative process such that

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma n_t, \tag{6.9}$$

   where n_t is the number of all relevant news for the company released at the day t, $\omega > 0$, $\alpha, \beta \geq 0$, $\alpha + \beta < 1$, γ, θ_1 and θ_2 are parameters of the model.

4. *GARCH model augmented with lagged news intensity.*

   Let r_t and r_t^* denote the log return of the company and log return of FTSE100 index at interval t respectively. We will consider a process $(\epsilon_t) = r_t - (\theta_1 + \theta_2 r_t^*)$ such that

$$\epsilon_t = \sigma_t u_t, t \in \mathbb{Z}, \tag{6.10}$$

   where (σ_t) is a nonnegative process such that

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma n_{t-1}, \tag{6.11}$$

   where n_{t-1} is the number of all relevant news for the company released at the day $t-1$, $\omega > 0$, $\alpha, \beta \geq 0$, $\alpha + \beta < 1$, γ, θ_1 and θ_2 are parameters of the model.

The first model is the model with contemporaneous trade volume. The second model is the model with lagged trade volume. The third model is the models with contemporaneous news intensity and the forth is with lagged news intensity. The number of news about a company at the day t is called the news intensity at the day t.

Empirical results can be found in the section 6.3. It will be shown that

- the GARCH(1,1) model augmented with volume V_t removes GARCH and ARCH effects for most of the FTSE100 companies.
- the GARCH(1,1) model augmented with lagged volume V_{t-1} does not remove GARCH and ARCH effects.
- the GARCH(1,1) model augmented with the news intensity n_t does not necessary remove GARCH and ARCH effects; however, using likelihood ratio test it will be shown that the model performs better than both the "pure" GARCH model for most of the FTSE100 companies.
- the GARCH(1,1) model augmented with the *lagged* news intensity n_{t-1} does not remove GARCH and ARCH effects.

6.2.2 Maximum likelihood estimation of augmented GARCH models

To calibrate the GARCH$(1,1)$ model we can use different methods including the least-square estimator or generalized moment method, but in this work we will apply the maximum likelihood approach.

One can generalized models (6.4)–(6.5), (6.6)–(6.7), (6.8)–(6.9) in the following way.

Let (ϵ_t) be a process such that

$$\epsilon_t = \sigma_t u_t, t \in \mathbb{Z}, \tag{6.12}$$

where (σ_t) is a nonnegative process such that

$$\sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + f(s_t, \mu), \tag{6.13}$$

where s_t is an exogenous time series, $f(\cdot, \mu) : \mathbb{R} \to \mathbb{R}$ is a continuous function with a vector of parameters μ, $\omega > 0$, $\alpha, \beta \geq 0$, $\alpha + \beta < 1$ and μ are parameters of the model.

The subsection describes quasi-maximum likelihood estimation (QML) of model (6.12), (6.13)

The vector of model parameters is

$$\theta = (\omega, \alpha, \beta, \mu)^T.$$

We will assume that θ belongs to the set

$$\Theta := \{(\omega, \alpha, \beta, \mu)^T : \omega > 0, \alpha > 0, \beta > 0 \in \mathbb{R}\}.$$

Denote

$$\theta^* = (\omega^*, \alpha^*, \beta^*, \mu^*)^T$$

the vector of the true values of parameters. The aim is to find θ^* that maximize a QML function given an observation sequence

$$\epsilon_0, \ldots, \epsilon_n$$

of length n.

Define the sequence $(\tilde{\sigma}_1, \ldots, \tilde{\sigma}_n)$ by recursion:

$$\tilde{\sigma}_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\tilde{\sigma}_{t-1}^2 + f(s_t, \mu),\ 1 \leq t \leq n,$$

where ϵ_0 and $\tilde{\sigma}_0$ are an initial values of ϵ's and σ's respectively.

Given the initial values, the Gaussian quasi-likelihood function can be written as follows

$$L_n(\theta) = L_n(\theta; \epsilon_1, \ldots, \epsilon_n) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi\tilde{\sigma}_t^2}} \exp\left(-\frac{\epsilon_t^2}{2\tilde{\sigma}_t^2}.\right)$$

The optimal estimation of θ is defined by

$$\tilde{\theta} = \arg\max_{\theta\in\Theta} L_n(\theta) = \arg\max_{\theta\in\Theta} F_n(\theta),$$

where

$$F_n(\theta) := -\sum_{t=1}^{n} \left(\frac{\epsilon_t^2}{\tilde{\sigma}_t^2} + \log\tilde{\sigma}_t^2\right)$$

is log quasi-likelihood function (constant terms are ignored).

6.3 Empirical Study

6.3.1 GARCH(1,1) Model with Volume

The aim of this section is to examine the impact of trading volume using GARCH approach suggested by Lamoureax and Lastrapes [1990].

1. Data description

Our sample covers a period ranging from July 5, 2005 to July 5, 2008 (i.e. 750 trading days). Our sample is composed of the 92 UK stocks that were part of the FTSE100 index in the beginning of 2005 and which survived through the period of 6 years. We have deleted 8 stocks. In this work we will present empirical results of only 5 company from the FTSE100. We focus our attention only on five companies traded on London Stock Exchange: AVIVA plc, BP, BT Group plc, Lloyd Banking Group, HSBC HLDG.

Daily stock closing prices (the last daily transaction price of the security), as well as daily transactions volume (number of shares traded during the day) are obtained from Yahoo Finance database. Results similar to one's presented in the chapter can be verified for all FTSE100 companies.

Table 6.2 provides preliminary descriptive statistics for the stock prices log returns and trading volumes.

Table 6.3 presents

- the list of stocks,
- the Kiefer-Salmon skewness test statistic (S)
- the Kiefer-Salmon kurtosis statistic (K)
- the Kiefer-Salmon joint statistic for normality (S+K)
- p-value of the Shapiro-Wilk statistic (marginal significance level)
- the Box-Ljung Q-statistic, constructed for maximum lag of 20.

Table 6.2: Descriptive statistics of five companies traded on London Stock Exchange

Company	T	Period	Mean	Std.Dev.	Skew.	Kurt.	Vol.Mean
Aviva	750	05/06/2005-05/06/2008	0.0001	0.0135	-0.3142	4.6735	1.12E+07
BP	750	05/06/2005-05/06/2008	0.0003	0.0121	0.0252	3.4093	7.76E+07
BT GROUP	750	05/06/2005-05/06/2008	0.0004	0.0134	0.4950	6.1042	4.46E+07
Lloyds	750	05/06/2005-05/06/2008	0.0000	0.0121	-0.2210	6.7680	3.26E+07
HSBC	750	05/06/2005-05/06/2008	-0.0001	0.0089	-0.0854	5.0921	4.13E+07

Table 6.3: Empirical properties of daily log returns and volumes for the five stocks in the sample

Company	S	K	SW(p)	Q(20)
AVIVA	0.0005	0.0000	0.9778	1292.7818
BP	0.7759	0.0376	0.9969	1730.2602
BT GROUP	0.0000	0.0000	0.9667	549.1024
Lloyds	0.0135	0.0000	0.9489	939.2284
HSBC	0.3361	0.0000	0.9746	1942.3059

It is well-known that S and K are $\chi^2(1)$-distributed, and $K + S$ is $\chi^2(2)$-distributed.

Based on the results presented in Table 6.3 one can conclude that the null hypothesis of normality is rejected for all stocks but BP.

The Box-Ljung Q-statistic tests for serial correlation in the daily volume series. It shows that there is no autocorrelation of log returns.

Consistent with the findings in Lamoureux and Lastrapes Lamoureax and Lastrapes [1990], we find that the p-values of Shapiro-Wilk statistic of log returns for all five companies are close to zero. We may conclude that all series are non-normal.

2. Empirical results

The GARCH model of Bollerslev [1986] provides a flexible and parsimonious approximation to conditional variance dynamics. Maximum likelihood estimates of "pure" GARCH(1,1) model (without Volume) for log returns of closing daily prices are presented in Table 6.4. Using GARCH estimates, Table 6.4 shows that the coefficients of the model are highly significant and volatility persistence, i.e. $\alpha + \beta$, is more than 0.9. It provides clear evidence of GARCH effect.

Table 6.4: Maximum likelihood estimates of "pure" GARCH(1,1) model (without Volume) for log returns of closing daily prices, $\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$

Company	α	β	$\alpha + \beta$	LLF_1
AVIVA	0.1209	0.8549	0.9758	2794.08
BP	0.0515	0.9208	0.9723	2867.78
BT GROUP	0.0770	0.8683	0.9454	2786.33
Lloyd	0.1230	0.8609	0.9839	2874.06
HSBC	0.1232	0.8568	0.9800	3112.39

The estimates of GARCH model with volume (6.4), (6.5) are presented in Table 6.5. The results show us that daily trading volume has significant explanatory power regarding the conditional volatility of daily log return for 3 of 5 companies (AVIVA, BP and BT Group). For Lloyds Group and HSBC Group there are not any changes in the level of persistence $\alpha + \beta$ compare to the results for the "pure" GARCH model.

The same contradictory picture is held for other FTSE100 companies. Once volume V_t is included as an explanatory variable in the equation, for many of FTSE100 companies the sum of $\alpha+\beta$ is significantly less than corresponding results in Table 6.4. One can see that once contemporaneous volume is included as an exogenous variable in the model, the impact of log return on volatility diminishes for most of FTSE100 companies.

Let us remind that $V_t = v_t/v^*$, where v_t is daily trade volume of a stock at the day t and $v^* = \max_t \; v_t$.

The estimates of γ in Table 6.5 for all five stocks are comparable with the square of unconditional standard deviation of the stocks (see Table 6.2).

Table 6.5: Maximum likelihood estimates of GARCH(1,1) model with Volume for log returns of the closing daily prices, $\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma V_t$

Company	α	β	γ	$\alpha+\beta$	LLF_2
AVIVA	0.2309	0.0012	1.22E-03	0.2321	2794.92
BP	0.1617	0.0000	7.19E-04	0.1617	2875.73
BT GROUP	0.1618	0.0000	9.38E-04	0.1618	2862.26
Lloyds	0.0928	0.8573	7.97E-05	0.9501	2883.04
HSBC	0.1387	0.8065	2.58E-05	0.9453	3118.81

To estimate the impact of lagged volume on volatility persistence in GARCH model, we consider a nonnegative process $\epsilon_t = \sigma_t u_t$, such that

$$\sigma_t = \omega + \alpha\epsilon_{t-1}^2 + \beta_1\sigma_{t-1} + \gamma V_{t-1}, \tag{6.14}$$

where V_{t-1} is the scaled trade volume for the company in day $t-1$, $\omega > 0$, $\alpha, \beta \geq 0$, $\alpha + \beta < 1$ and γ are parameters of the model.

The results presented in Table 6.12 show that there are no evidence of vanishing effect of log return on volatility. Moreover, estimates of parameters α, β are close to corresponding ones in Table 6.4.

Table 6.6: Maximum likelihood estimates of GARCH(1,1) model with *Lagged* Volume for log returns of the closing daily prices, $\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma V_{t-1}$

Company	α	β	γ	$\alpha+\beta$	LLF_3
AVIVA	0.1150	0.8683	3.55E-05	0.9832	2794.78
BP	0.0516	0.9208	3.41E-11	0.9723	2867.78
BT GROUP	0.0761	0.8803	2.40E-11	0.9564	2867.22
Lloyds	0.0514	0.9208	3.32E-11	0.9721	2790.64
HSBC	0.1181	0.8574	1.20E-05	0.9755	3113.12

Note that the GARCH model (the null model) is a special case of the GARCH model augmented with volume (the alternative model). Therefore, to compare the fit of two models it can be used a likelihood ratio test (see e.g. Cox and Hinkley [1974]). It is the most common approach to testing problem. The test was introduced by Neyman and Pearson in 1928. It compares the maximum likelihood under the alternatives with that under the hypothesis. Let us remind the main idea of the test. Let the likelihood function of θ is

$$LF(x, \theta) = p_\theta(x),$$

i.e. the probability density (or probability) of x considered as a function of θ. It is widely considered a (relative) measure of support that the observation x gives to the parameter θ. Then the likelihood ratio compares the best explanation the data provide for the alternatives with the best explanations for the hypothesis. The likelihood ratio is a function of the data x, therefore it is a statistic. The likelihood-ratio test rejects the null hypothesis if the value of this statistic is too small. We must compare the value of likelihood ratio to a critical value to decide whether to reject the null model in favor of the alternative model.

Results of likelihood ratio test for GARCH model (null model) and GARCH model augmented with volume (alternative model) one can find in Table 6.7. For four of five companies the alternative model is preferable with confidence level of 1%.

Table 6.7: Results of the likelihood ratio test for GARCH model and GARCH model augmented with volume

Company	LLF_1	LLF_2	$2(LLF_2 - LLF_1)$	$\chi^2(1)$, 1%	Null Hyp.
AVIVA	2794.08	2794.92	1.70	6.64	accepted
BP	2867.78	2875.73	15.89	6.64	rejected
BT Group	2786.33	2862.26	151.87	6.64	rejected
Lloyds	2874.06	2883.04	17.96	6.64	rejected
HSBC	3112.39	3118.81	12.84	6.64	rejected

Results of likelihood ratio test for GARCH model (null model) and GARCH model augmented with *lagged* volume (alternative model) one can find in Table 6.8. For four of five companies the alternative model is rejected with confidence level of 1%.

Figure 6.1 presents the volatility forecast of GARCH model and GARCH model with Volume for HSBC stock market closing daily prices.

Table 6.8: Results of the likelihood ratio test for GARCH model and GARCH model augmented with *lagged* volume

Company	LLF_1	LLF_3	$2(LLF_3 - LLF_1)$	$\chi^2(1)$, 1%	Null Hyp.
AVIVA	2794.08	2794.78	1.42	6.64	accepted
BP	2867.78	2867.78	0.00	6.64	accepted
BT Group	2786.33	2790.64	8.63	6.64	rejected
Lloyds	2874.06	2874.64	1.15	6.64	accepted
HSBC	3112.39	3113.12	1.45	6.64	accepted

6.3.2 Monte Carlo simulation for likelihood ratio statistic

One of possible ways of testing one model against another one is to form the likelihood ratio statistic. This test has been discussed in the papers Lee and Brorsen [1997] and Kim et al. [1998]. In this subsection we use this approach to test the GARCH-volume model against GARCH model (as well as the GARCH-lagged-volume model against GARCH model).

Let H_0 denote the GARCH-volume model and H_1 denote the GARCH model. Let ϵ_t be a random variable that has a mean and a variance conditionally on the information set I_{t-1}.

Denote the corresponding log likelihood functions by $LLF_{H_0}(\epsilon;\theta_0)$ and $LLF_{H_1}(\epsilon;\theta_1)$, respectively.

We will consider the test statistic defined by

$$LR = 2(LLF_{H_0}(\epsilon;\widetilde{\theta}_0) - LLF_{H_1}(\epsilon;\widetilde{\theta}_1)). \tag{6.15}$$

While the asymptotic null distribution of (6.15) is unknown, it can be approximated by Monte Carlo simulation.

We can assume that the GARCH-volume model is the null model and that $\widetilde{\theta}_0$ is the true parameter. Using Monte Carlo approach we will generate N realizations of T observations

$$\epsilon^{(i)} = (\epsilon_t^{(i)})_{t=1}^T, \;\; i = 1, \ldots, N,$$

from this model. Then we will estimate both models and calculates the value of (6.15) using each realization $\epsilon^{(i)}$.

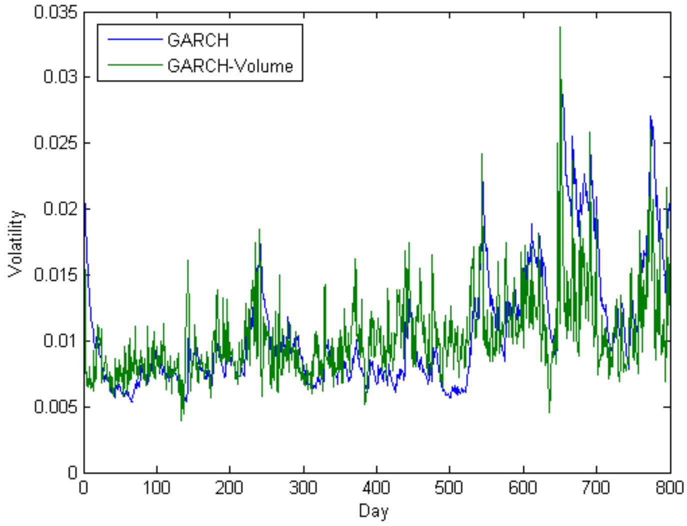

Figure 6.1: Volatility forecast of GARCH model and GARCH model with Volume for HSBC stock market closing daily prices (05/06/2005-05/06/2008)

Ranking the N values gives an empirical distribution with which one compares the original value of (6.15). The true value of $\widetilde{\theta}_0$ is unknown, but the approximation error due to the use of $\widetilde{\theta}_0$ as a replacement vanishes asymptotically as $T \to \infty$.

If the value of (6.15) is more or equal to the $100(1-\alpha)\%$ quantile of the empirical distribution, the null model is rejected at significance level α. As it was mentioned in Lee and Brorsen [1997] the models under comparison need not have the same number of parameters, and the value of the statistic can also be negative. Reversing the roles of the models, it can be possible to test GARCH models against GARCH-volume model.

We present the results of the Monte Carlo simulation for the likelihood ratio statistic to compare

1. the GARCH model and the GARCH model with volume,
2. the GARCH model and the GARCH model with lagged volume

on the finite sample performance of the MLE estimator. In particular, we study the significance of the MLE estimators of the parameters of the variance equation (equations (6.5) and (6.7)).

1. Monte Carlo test results for GARCH model with volume

The data-generating model is defined by equations (6.4) and equation (6.5) given before. Notice that the error term in the mean equation is drawn from a normal distribution with mean zero and variance that changes over time according to equation (6.5).

Finally, we have set the number of trials N in each Monte Carlo experiment to 500.

MATLAB code is presented in Appendix B. The experiment took more than 10 hours on standard PC.

- Function MC_Volume_Sim_Head.m
  1. loads two files with input data (log returns and volumes);
  2. runs function MC_GARCH_Volume_simulation.m for $N = 500$ times;
  3. outputs of the function MC_Volume_sim_head.m are mean and variance of LLF ratio statistic, as well as means and variances of all parameters of GARCH model with volume ($\omega, \alpha, \beta, \gamma$).
- Function MC_GARCH_Volume_simulation.m is a function with input parameters
  1. arrays of input data (log returns and volumes);
  2. initial values of model parameters.

  Function MC_GARCH_Volume_simulation.m simulates the sequence of ϵ's. Based on Equation (6.5) it uses iteratively MATLAB function random('normal',0,1) to get a sequence of ϵ's.

The results of the Monte Carlo simulation are presented below.

Figures 6.2 and 6.3 illustrate the empirical distribution of parameters α and β of GARCH model with Volume for HSBC, respectively.

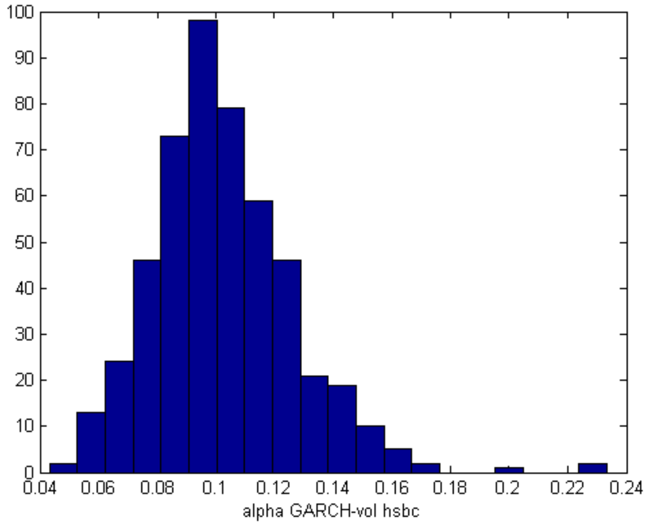

Figure 6.2: Histogram of parameter α of GARCH model with Volume for HSBC (Monte Carlo simulation runs $N = 500$ times)

Results of likelihood ratio test for GARCH model and GARCH model augmented with volume one can find in Table 6.9. For four of five companies the alternative model H_1 is rejected with confidence level of 10%. Moreover, the Monte Carlo simulation shows that almost all parameters of GARCH-volume model are significant at least with confidence level of 25%.

2. Monte Carlo test results for GARCH model with lagged volume

Now we present the results of the Monte Carlo simulation for the likelihood ratio statistic to compare GARCH model and GARCH-volume-lagged model. In particular, we study the significance of the MLE estimators of the parameters of the variance equation (equations (6.7)).

The data-generating model (i.e. GARCH model with lagged volume) is defined by equations (6.6) and equation (6.7) given before. Notice that the error term in the mean equation is drawn from a normal distribution with mean zero and variance that changes over time according to equation (6.7).

Finally, we have set the number of trials N in each Monte Carlo experiment to 500. Given that not always the program that solves the model numerically will achieve convergence, the final number of (valid) trials is less than 500.

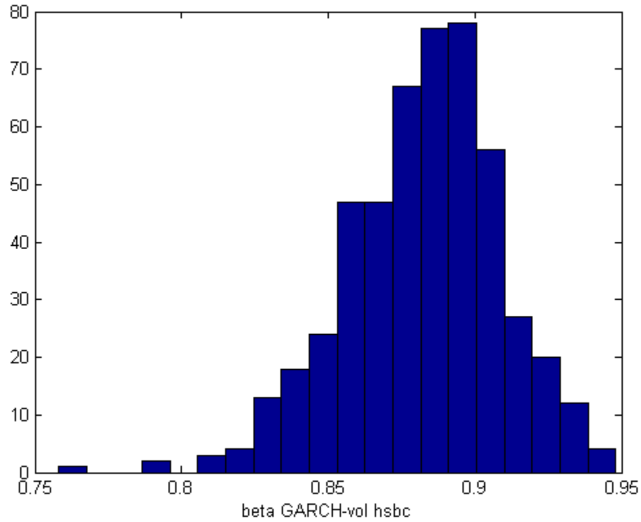

Figure 6.3: Histogram of parameter β of GARCH model with Volume for HSBC (Monte Carlo simulation runs $N = 500$ times)

This happens particularly in the GARCH model with lagged volume.

MATLAB code is presented in Appendix.

The results of the Monte Carlo simulation are presented below. Results of likelihood ratio test for GARCH model augmented with *lagged* volume against GARCH model are presented in Table 6.10. For all of five companies the GARCH model augmented with *lagged* volume do not perform better than alternative model H_1 (GARCH model) with confidence level of 10%. Moreover, the Monte Carlo simulation shows that almost all parameters of GARCH-volume model are significant at least with confidence level of 25%.

Figures 6.4 and 6.5 present histograms of parameters α and β of GARCH model with Volume for HSBC, respectively.

Table 6.9: Monte Carlo results for GARCH(1,1) model with Volume. Values in parenthesis are standard deviations. The symbols ˇ, ˆ and $*$, indicate if the parameter estimation is significant with levels of 5%, 10% and 25% respectively.

Company	ω	α	β	γ	LLF Ratio
AVIVA	2.81E-05*	0.1920	0.3186ˆ	2.02E-3*	3.5217
	(1.70E-05)	(0.2888)	(0.1715)	(1.77E-3)	(1.1906)
BP	1.58E-05*	0.0694ˇ	0.2809	1.87E-4*	9.1958*
	(1.04E-05)	(0.0282)	(0.3265)	(1.20E-4)	(2.3943)
BT GROUP	8.75E-06ˇ	0.0421*	0.1309	4.69E-4ˇ	8.5014*
	(1.84E-05)	(0.0294)	(0.2086)	(1.66E-4)	(3.6952)
Lloyds	1.56E-07ˆ	0.1506ˇ	0.8505ˇ	1.57E-5ˇ	13.9909*
	(6.44E-08)	(0.0358)	(0.0337)	(7.37E-6)	(6.0937)
HSBC	2.14E-07ˆ	0.1024ˇ	0.8822ˇ	1.14E-5ˇ	10.0327*
	(5.18E-08)	(0.0242)	(0.0268)	(4.12E-6)	(3.0207)

6.3.3 GARCH(1,1) Model Augmented with News Intensity

1. Data Description

Our sample covers a period ranging from July 5, 2005 to July 5, 2008 (i.e. 750 trading days). Our sample is composed of the 92 UK stocks that were part of the FTSE100 index in the beginning of 2005 and which survived through the period of 6 years. We have deleted 8 stocks. In this work we will present empirical results of only 5 company from the FTSE100. We focus our attention only on five companies traded on London Stock Exchange: AVIVA, BP, BT Group, Lloyd Banking Group, HSBC. Daily stock closing prices (the last daily transaction price of the security) are obtained from Yahoo Finance database. Table 6.2 provides preliminary descriptive statistics for the stock prices log returns and trading volumes.

All news analytics data were given by Raven Pack News Analytics (RPNA). RPNA is a news sentiment analysis service that provides a look into the sentiment of more than 28,000 publicly traded companies worldwide. Each score is a weighed balance of sentiment in articles published by professional newswires (such as Dow Jones and Reuters) and hundreds of financial sites, online newspapers and even blogs.

Table 6.10: Monte Carlo results for GARCH(1,1) model with *Lagged* Volume. Values in parenthesis are standard deviations. The symbols ˘, ˆ and $*$, indicate if the parameter estimation is significant with levels of 5%, 10% and 25% respectively.

Company	ω	α	β	γ	LLF Ratio
aviva	1.97E-06	0.1408*	0.8560*	4.87621E-5ˆ	1.3769
	(2.63E-06)	(0.0306)	(0.0299)	(2.92E-5)	(1.6562)
BP	7.07E-06˘	0.0838*	0.8902*	7.25E-11	0.0144
	(2.31E-06)	(0.0164)	(0.0221)	(7.24E-11)	(1.0436)
BT	1.08E-05*	0.1050˘	0.8651˘	9.25E-11˘	0.0921
	(5.02E-06)	(0.0289)	(0.0387)	(1.46E-11)	(0.8643)
lloyd	3.34E-06ˆ	0.1017˘	0.8970˘	9.75E-11˘	1.0327
	(1.30E-06)	(0.0190)	(0.0180)	(1.51E-11)	(1.0486)
hsbc	3.57E-07	0.1405˘	0.8565˘	7.19E-5*	3.9909
	(6.52E-07)	(0.0331)	(0.0311)	(3.02E-5)	(2.6708)

For each news wire, we have got the following fields (Table 6.1): time stamp, company name, company id, relevance of the news, event category, event sentiment, novelty of the news, novelty id, composite sentiment score of the news, word/phrase level score, projections by company, editorials & commentary, reports corp actions, news impact projection, story ID. Company, relevance score, composite sentiment score are the main fields of interest. One piece of news can of course concern several companies, industries and subjects. To avoid any redundancy and duplicate announcements that do not bring any additional information value, we restrict the sample to news released with high relevance score (more or equal to 90). We do not eliminate all news releases with the same headlines and lead paragraphs, since it is supposed that the number of the same news published by different news agencies reflects the importance of the news.

For example, there was more than 8000 financial BP news releases with relevance ≥ 90 over the whole sample period.

Figure 6.6 shows the evolution over time of the total daily number of news wires for BP company.

Figure 6.7 displays the evolution over time of the rolling mean of the number of news wires with 5-days window.

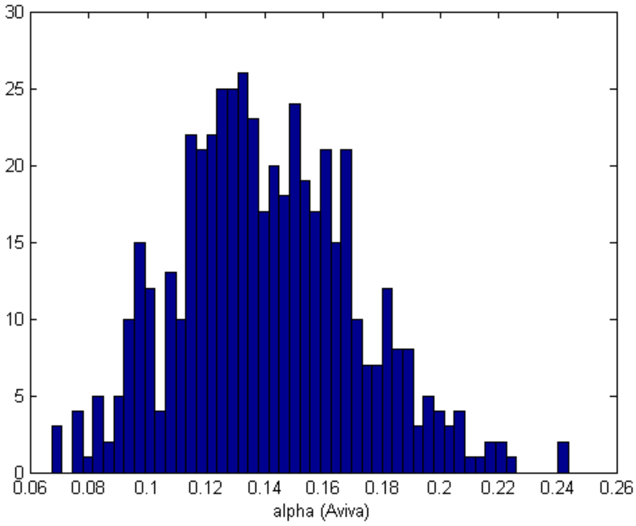

Figure 6.4: Histogram of parameter α of GARCH model with lagged Volume for AVIVA company (Monte Carlo simulation runs $N = 500$ times)

One can see that there is no any clear trend both in Figure 6.6 and Figure 6.7. It could indicate that the news time-series is rather stationary and reduce the risk of dummy results due to a possible simultaneous increase over time of the stock volatility. Some periods have rate of news intensity below the average (e.g. holidays, Christmas time). On the other hand, one can witnessed the increase of the rate at the periods of the quarterly reports and releases of the intermediate figures and earnings of companies.

One can see a clear presence of weekly seasonality in the data. For example, Figure 6.8 shows that the average number of British Petroleum' news announcements released during the week-end is much lower than the one of the other weekdays. The same picture is held for all FTSE100 companies indeed. Since that we exclude all weekend news from our analysis.

Figure 6.9 shows the evolution of the frequency of announcements arrivals throughout the day (at New-York's time) for BP company. There are picks at 9 am and 4 pm, and the activity seems to be globally more sustained in the morning than in the afternoon. The "lunch drop" is easily recognizable.

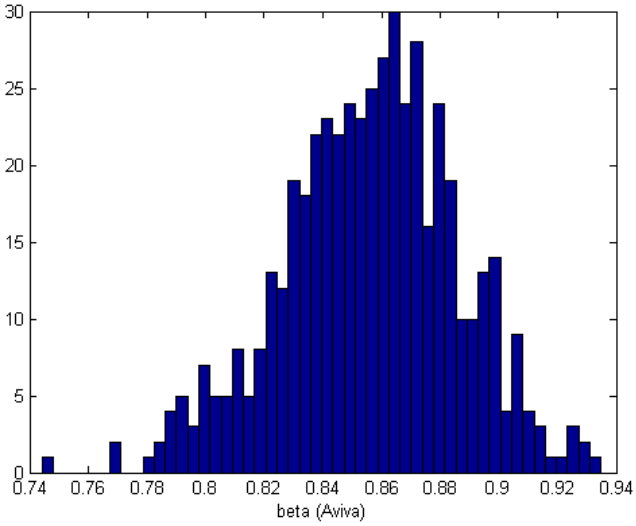

Figure 6.5: Histogram of parameter β of GARCH model with lagged Volume for AVIVA company (Monte Carlo simulation runs $N = 500$ times)

2. Empirical Results

The estimates of GARCH model with news intensity (6.8), (6.9) are presented in Table 6.11. The results show us that daily news intensity has some explanatory power regarding the conditional volatility of daily log return. Once news intensity n_t is included as an explanatory variable in the equation, the sum of $\alpha + \beta$ is less than corresponding results in Table 6.4. One can see that once contemporaneous news intensity is included as an exogenous variable in the model, the GARCH effect slightly diminishes for some companies (BP and HSBC Group).

To estimate the impact of lagged news intensity on volatility persistence in GARCH model, we consider a nonnegative process $\epsilon_t = \sigma_t u_t$, such that

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2 + \gamma n_{t-1}, \tag{6.16}$$

where n_{t-1} is the news intensity for the company in day $t-1$, $\omega > 0$, $\alpha, \beta \geq 0$, $\alpha + \beta < 1$ and γ are parameters of the model.

The results presented in Table 6.12 show that there are no evidence of vanishing GARCH effect. Moreover, estimates of parameters α, β are close to corresponding ones in Table 6.4.

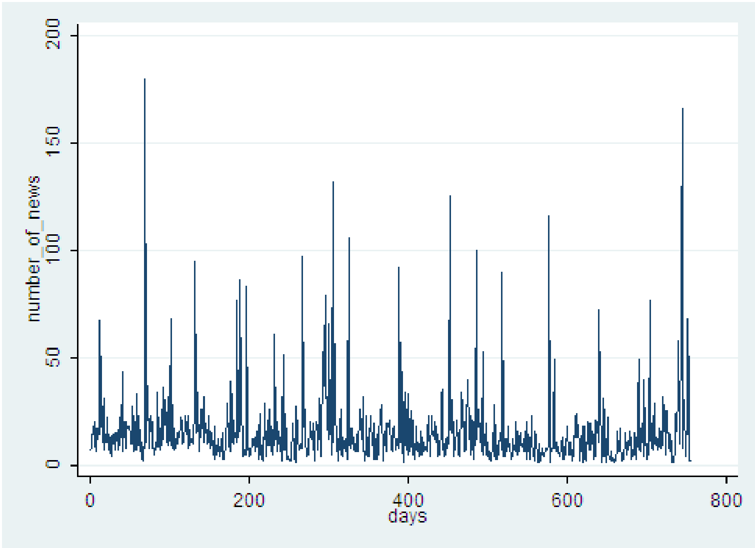

Figure 6.6: The dynamics of British Petroleum′ news announcements

Results of likelihood ratio test for GARCH model (null model) and GARCH model augmented with news intensity (alternative model) one can find in Table 6.13. For all five companies the alternative model is preferable with confidence level of 1%.

Results of likelihood ratio test for GARCH model (null model) and GARCH model augmented with *lagged* news intensity (alternative model) one can find in Table 6.14. For all five companies the alternative model is rejected with confidence level of 1%.

Figure 6.10 presents the volatility forecast performed by GARCH model and GARCH model with Volume for HSBC stock market closing daily prices.

6.3.4 Monte Carlo simulation for GARCH model with news intensity

We will use the test proposed in Lee and Brorsen [1997] and Kim et al. [1998]. The latter authors suggested it for testing the GARCH model against the autoregressive stochastic volatility model or vice versa.

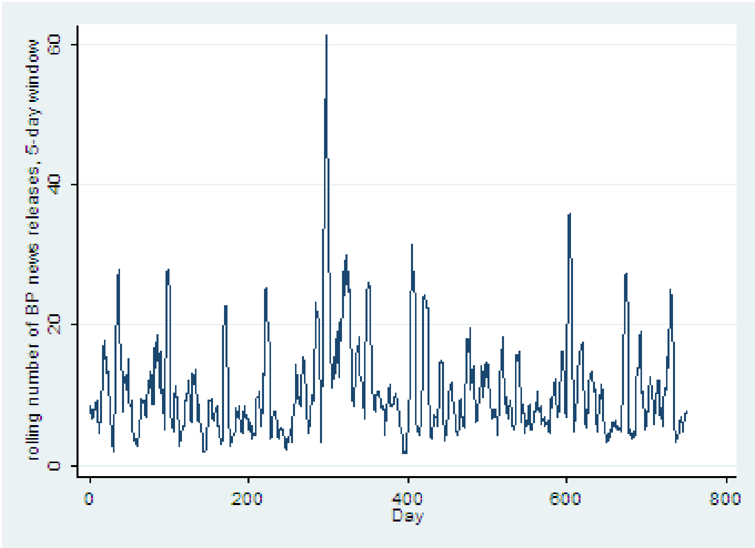

Figure 6.7: The rolling mean of the number of British Petroleum' news announcements, 5-days window (January 3, 2005 – December 31, 2007)

Let H_0 be the GARCH-news model and H_1 the GARCH one. The test is based on the log likelihood ratio

$$LR = 2(LLF_{H_0}(\epsilon; \widetilde{\theta}_0) - LLF_{H_1}(\epsilon; \widetilde{\theta}_0)).$$

where $LLF_{H_1}(\epsilon; \widetilde{\theta}_1)$ and $LLF_{H_0}(\epsilon; \widetilde{\theta}_0)$ are the maximized likelihood function under the GARCH model and under the GARCH-news model, respectively.

The asymptotic distribution of LR is unknown. Therefore we can use an empirical distribution constructed by simulation. We use the LR as test statistic in a Monte Carlo hypothesis test. Unfortunately the true parameters are unknown. Here, only a consistent estimate of the parameters is available. When the parameters must be estimated, Monte Carlo hypothesis tests are no longer exact, but still asymptotically valid; see Kim et al. [1998].

Note that in Chen and Kuan [2002] proposed yet another method based on the pseudo- score, whose estimator under the null hypothesis and assuming the customary regularity conditions is asymptotically normally distributed. This result forms the basis for a χ^2-distributed test statistic (see Chen and Kuan [2002] for details).

Results of small-sample simulations in Malmsten [2004] indicate that the

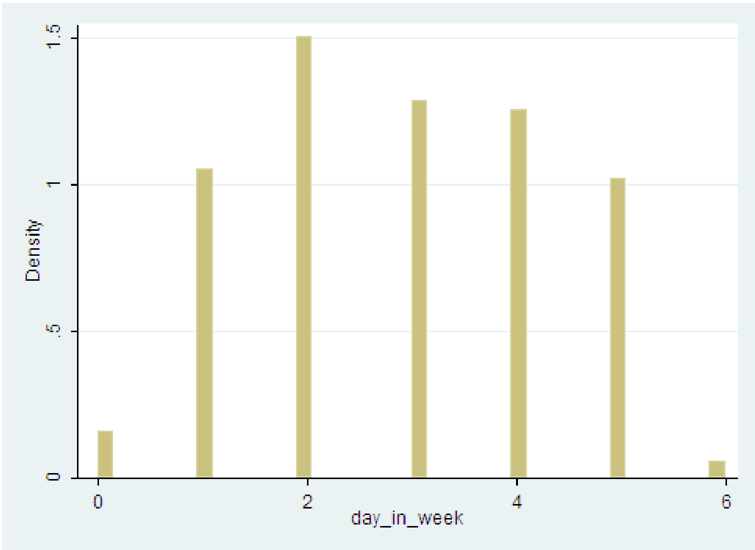

Figure 6.8: The average number of British Petroleum' news announcements released per day in week

pseudo-score test tends to be oversized. Furthermore, the Monte Carlo likelihood ratio statistic seems to have consistently higher power than the encompassing test, which suggests that the former rather than the latter should be applied in practice.

The data-generating model is defined by equations (6.8) and equation (6.9) given before. Notice that the error term in the mean equation is drawn from a normal distribution with mean zero and variance that changes over time according to equation (6.9).

We present the results of the Monte Carlo simulation for the likelihood ratio statistic to compare

1. the GARCH model and the GARCH model with news intensity,
2. the GARCH model and the GARCH model with lagged news intensity

on the finite sample performance of the MLE estimator. In particular, we study the significance of the MLE estimators of the parameters of the variance equation (equations (6.9) and (6.11)).

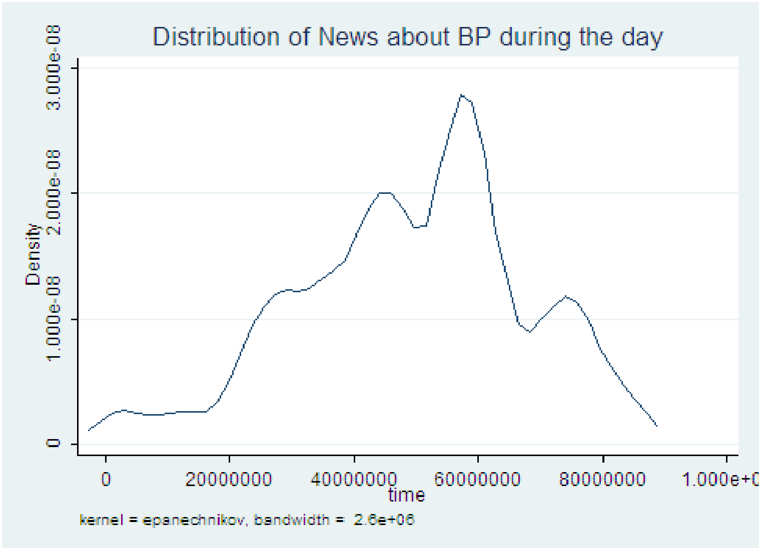

Figure 6.9: The average number of British Petroleum′ news announcements released during the day

1. Monte Carlo test results for GARCH model with news intensity

The data-generating model is defined by equations (6.8) and equation (6.9) given before. Notice that the error term in the mean equation is drawn from a normal distribution with mean zero and variance that changes over time according to equation (6.9).

Finally, we have set the number of trials N in each Monte Carlo experiment to 500. Given that not always the program that solves the model numerically will achieve convergence, the final number of (valid) trials is less than 500. This happens particularly in the GARCH-volume model.

MATLAB code is presented in Appendix B. The experiment took more than 14 hours on standard PC.

- Function MC_News_sim_head.m

  1. loads files with input data (arrays of log returns of the company stocks, log returns of of FTSE100 index, arrays of numbers of news per day);
  2. runs function MC_GARCH_News_simulation.m for $N = 500$ times;

Table 6.11: Maximum likelihood estimates of GARCH(1,1) model with news intensity for log returns of the closing daily prices, $(\epsilon_t) = r_t - (\theta_1 + \theta_2 r_t^*)$, $\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma n_t$

Company	α	β	γ	θ_1	θ_2	$\alpha+\beta$	LLF_4
AVIVA	0.0001	0.9562	2.24E-06	-4.36E-04	1.2086	0.9563	2963.51
BP	0.0010	0.1428	5.39E-06	-1.17E-04	0.9814	0.1437	3086.74
BT GROUP	0.0811	0.7595	8.36E-06	1.23E-05	0.7568	0.8406	2944.74
Lloyds	0.1500	0.8109	4.39E-06	7.38E-05	0.9655	0.9609	3134.81
HSBC	0.1448	0.1343	3.68E-06	-4.16E-04	0.8534	0.2791	3309.99

Table 6.12: Maximum likelihood estimates of GARCH(1,1) model with *Lagged* news intensity for log returns of the closing daily prices, $(\epsilon_t) = r_t - (\theta_1 + \theta_2 r_t^*)$, $\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma n_{t-1}$

Company	α	β	γ	θ_1	θ_2	$\alpha+\beta$	LLF_5
AVIVA	0.1178	0.8076	5.80E-12	-5.51E-04	1.1715	0.9254	3039.36
BP	0.0199	0.9494	2.94E-13	1.25E-04	0.9325	0.9694	3113.03
BT GROUP	0.0863	0.7105	1.12E-20	1.43E-05	0.8129	0.7968	2907.11
Lloyds	0.1366	0.8380	1.76E-14	3.51E-05	0.9716	0.9746	3113.29
HSBC	0.1544	0.7830	3.08E-24	-1.14E-05	0.7704	0.9375	3306.47

3. outputs of the function MC_News_sim_head.m are mean and variance of LLF ratio statistic, as well as means and variances of all parameters of GARCH model with news intensity $(\omega, \alpha, \beta, \gamma)$.

- Function MC_GARCH_News_simulation.m is a function with input parameters

  1. arrays of input data (arrays of log returns of the company stocks, log returns of of FTSE100 index, arrays of numbers of news per day);
  2. initial values of model parameters.

  Function MC_GARCH_Volume_simulation.m simulates the sequence of ϵ's. Based on Equation (6.9) it uses iteratively MATLAB function random('normal',0,1) to get a sequence of ϵ's.

The results of the Monte Carlo simulation are presented below.

Table 6.13: Results of the likelihood ratio test for GARCH model and GARCH model augmented with news intensity

Company	LLF_1	LLF_4	$2(LLF_4 - LLF_1)$	$\chi^2(3), 1\%$	Null Hyp.
AVIVA	2794.08	2963.51	338.88	7.82	rejected
BP	2867.78	3086.74	437.94	7.82	rejected
BT Group	2786.33	2944.74	316.83	7.82	rejected
Lloyds	2874.06	3134.81	521.51	7.82	rejected
HSBC	3112.39	3309.99	395.21	7.82	rejected

Table 6.14: Results of the likelihood ratio test for GARCH model and GARCH model augmented with *lagged* news intensity

Company	LLF_1	LLF_5	$2(LLF_5 - LLF_1)$	$\chi^2(3), 1\%$	Null Hyp.
AVIVA	2794.08	3039.36	490.57	7.82	rejected
BP	2867.78	3113.03	490.50	7.82	rejected
BT Group	2786.33	2907.11	241.57	7.82	rejected
Lloyds	2874.06	3113.29	478.46	7.82	rejected
HSBC	3112.39	3306.47	388.17	7.82	rejected

Figures 6.11 present the histogram of LL ratio for HSBC.

Results of likelihood ratio test for GARCH model and GARCH model augmented with news intensity one can find in Table 6.15. For four of five companies the alternative model H_1 is rejected with confidence level of 10%. Moreover, the Monte Carlo simulation shows that almost all parameters of GARCH-news model are significant at least with confidence level of 25%.

2. Monte Carlo test results for GARCH model with lagged news intensity

Now we present the results of the Monte Carlo simulation for the likelihood ratio statistic to compare GARCH model and GARCH-lagged-news model. In particular, we study the significance of the MLE estimators of the parameters of the variance equation (equations (6.7)).

The data-generating model (i.e. GARCH model with lagged volume) is defined by equations (6.6) and equation (6.7) given before. Notice that the error term in the mean equation is drawn from a normal distribution with mean zero and variance that changes over time according to equation (6.7).

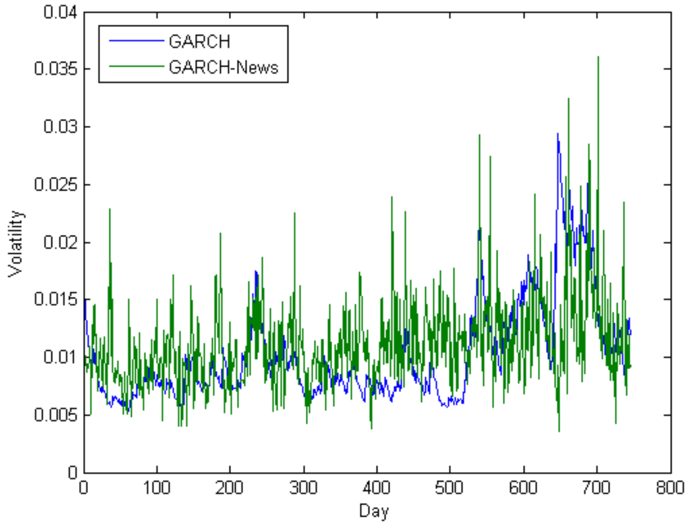

Figure 6.10: Volatility forecast by GARCH model and GARCH model with news intensity for HSBC stock market closing daily prices (05/06/2005-05/06/2008)

Finally, we have set the number of trials N in each Monte Carlo experiment to 500. MATLAB code is presented in Appendix B. The computations took more than 14 hours on standard PC.

The results of the Monte Carlo simulation are presented below. Results of likelihood ratio test for GARCH model and GARCH model augmented with *lagged* news intensity one can find in Table 6.16. For all of five companies the alternative model H_1 (GARCH model) is not rejected with confidence level of 10%. However, the Monte Carlo simulation shows that almost all parameters of GARCH-volume model are significant at least with confidence level of 10%. Figures 6.12 present the histogram of LL ratio for HSBC.

3. Conclusion

The results show us that daily trading volume does have significant explanatory power regarding the conditional volatility of daily log return. Based on empirical study for stocks of some of FTSE100 companies we may conclude that once contemporaneous volume is included as an exogenous

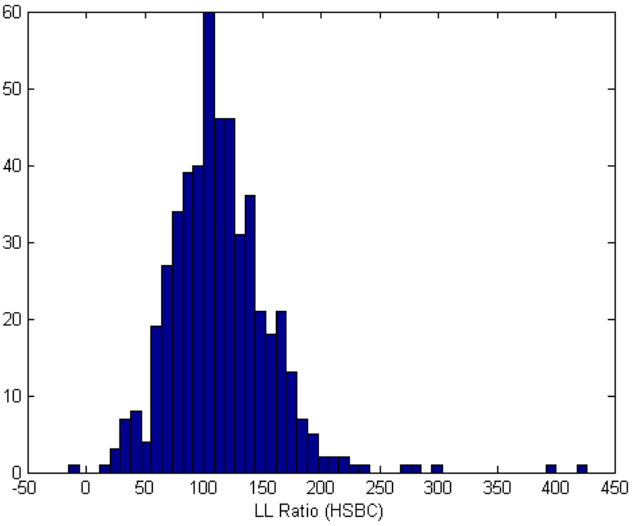

Figure 6.11: Histogram of log likelihood ratio for GARCH model with news intensity for HSBC (Monte Carlo simulation runs $N = 500$ times)

variable in the model, the GARCH effect diminishes for most of the FTSE100 companies.

However, the results presented in Table 6.12 show that there are no evidence of vanishing GARCH effect for the GARCH(1,1) model augmented with *lagged* volume V_{t-1}.

Then we merely reproduce the methodology commonly used in the literature (see Kalev et al. [2004] and Janssen [2004]). It sums up to insert the number of daily announcements concerning a stock into the equation of its variance in a GARCH (1,1) model. The estimation of the latter after inclusion of the variable NBN (daily number of news releases) converges for all 5 stocks. It was shown that the GARCH(1,1) model augmented with the news intensity n_t (the number of daily announcements) does not necessarily remove GARCH and ARCH effects.

However, the likelihood ratio test shows that the GARCH(1,1) model augmented with the news intensity performs better than both the GARCH model.

Then we use the Monte Carlo simulation of likelihood ratio test discussed

Table 6.15: Monte Carlo results for GARCH(1,1) model with news intensity. Values in parenthesis are standard deviations. The symbols ˇ, ˆ and $*$, indicate if the parameter estimation is significant with levels of 5%, 10% and 25% respectively.

Company	ω	α	β	LLF Ratio
AVIVA	6.87E-06ˇ	0.1209ˇ	0.8549ˇ	7.5696
	(2.79E-7)	(7.22E-4)	(9.33E-5)	(15.9179)
BP	5.11E-06ˇ	0.0515ˇ	0.9208ˇ	70.8573ˇ
	(9.32E-8)	(5.07E-3)	(2.33E-2)	(14.2960)
BT Group	1.32E-05ˆ	0.0770ˇ	0.8683ˇ	83.5708ˆ
	(8.98E-7)	(4.02E-3)	(7.55E-7)	(28.7227)
Lloyds	4.91E-06ˇ	0.1230ˇ	0.8609ˇ	25.1883*
	(5.93E-8)	(9.44E-5)	(2.33E-7)	(13.6620)
HSBC	2.72E-06ˇ	0.1232ˇ	0.8568ˇ	114.4146ˆ
	(2.96758E-10)	(1.1669E-6)	(7.33481E-5)	(43.9317)

in the papers Lee and Brorsen [1997] and Kim et al. [1998]. This approach has been used for testing the augmented GARCH models against GARCH model. In particular, we study the significance of the MLE estimators of the parameters of the variance equation

Results of likelihood ratio test for GARCH model and GARCH model augmented with volume show that for four of five companies the alternative model (GARCH model) is rejected with confidence level of 10%. Moreover, the Monte Carlo simulation shows that almost all parameters of GARCH-volume model are significant at least with confidence level of 25%.

Then we have presented the results of the Monte Carlo simulation for the likelihood ratio statistic to compare GARCH model and GARCH-volume-lagged model. The results of the Monte Carlo simulation show that for all of five companies the GARCH model augmented with *lagged* volume do not perform better than alternative model (GARCH model) with confidence level of 10%. However, the Monte Carlo simulation shows that almost all parameters of GARCH-volume model are significantly different from zero at least with confidence level of 25%.

Results of likelihood ratio test for GARCH model and GARCH model augmented with news intensity show that for four of five companies the alternative model (GARCH model) is rejected with confidence level of 10%.

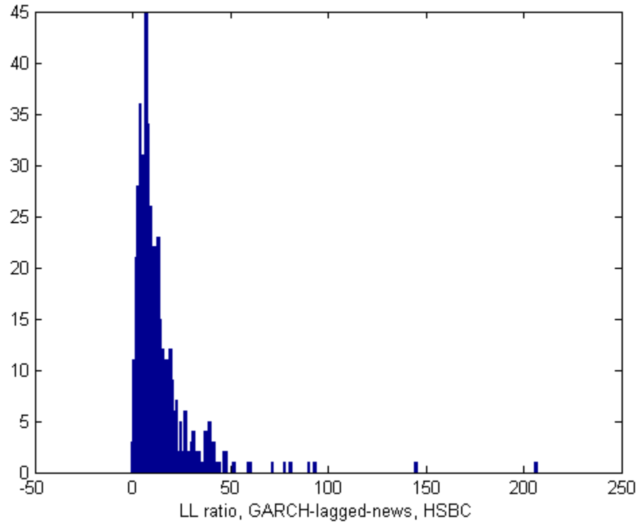

Figure 6.12: Histogram of log likelihood ratio for GARCH model with news intensity for HSBC (Monte Carlo simulation runs $N = 500$ times)

Moreover, the Monte Carlo simulation shows that almost all parameters of GARCH-news model are significant at least with confidence level of 25%.

Monte Carlo simulation of likelihood ratio test for GARCH model and GARCH model augmented with news intensity show that for four of five companies the alternative model H_1 is rejected with confidence level of 10%. Moreover, the Monte Carlo simulation shows that almost all parameters of GARCH-news model are significant at least with confidence level of 25%.

The results of the Monte Carlo simulation for GARCH model and GARCH model augmented with *lagged* news intensity show that we can not reject the alternative model H_1 (GARCH model) with confidence level of 10% for all five companies. However, the Monte Carlo simulation shows that almost all parameters of GARCH-volume model are significant at least with confidence level of 10%.

Table 6.16: Monte Carlo results for GARCH(1,1) model with *Lagged* news intensity. Values in parenthesis are standard deviations. The symbols ˇ, ˆ and $*$, indicate if the parameter estimation is significant with levels of 5%, 10% and 25% respectively.

Company	ω	α	β	γ	LLF Ratio
AVIVA	6.87E-06ˇ	0.1209ˇ	0.8549ˇ	6.87E-06ˇ	11.71714179
	(2.80E-20)	(7.22E-16)	(9.34E-15)	(2.80E-20)	(1.46E+01)
BP	5.11E-06ˇ	0.0515ˇ	0.9208ˇ	5.11E-06ˇ	31.62763573
	(9.33E-21)	(5.07E-16)	(2.33E-15)	(9.33E-21)	(7.87E+00)
BT Group	1.32E-05ˇ	0.077ˇ	0.8683ˇ	1.32E-05ˇ	14.98114309
	(8.99E-20)	(4.03E-16)	(7.56E-15)	(8.99E-20)	(1.16E+01)
Lloyds	4.91E-06ˇ	0.123ˇ	0.8609ˇ	4.91E-06ˇ	22.65617058
	(5.94E-20)	(9.45E-16)	(2.33E-15)	(5.94E-20)	(2.43E+01)
HSBC	2.72E-06ˇ	0.1232ˇ	0.8568ˇ	2.72E-06ˇ	13.42066052
	(2.97E-20)	(1.17E-15)	(7.33E-15)	(2.97E-20)	(1.61E+01)

Chapter 7

GARCH models with Jumps

7.1 GARCH model with Jumps

7.1.1 Model Description

GARCH-Jump model was proposed and studied in Maheu and McCurdy [2004]. This paper proposes a model of conditional variance of returns implied by the impact of different type of news.

Let X_t be the log return of a particular stock or the market portfolio from time $t-1$ to time t. Let I_{t-1} denote the past information set containing the realized values of all relevant variables up to time $t-1$. Suppose investors know the information in I_{t-1} when they make their investment decision at time $t-1$. Then the relevant expected return μ_t to the investors is the conditional expected value of X_t, given I_{t-1}, i.e.

$$\mu_t = E(X_t|I_{t-1}).$$

The relevant expected volatility σ_t^2 to the investors is conditional variance of X_t, given I_{t-1}, i.e.

$$\sigma_t^2 = Var(X_t|I_{t-1}).$$

Then

$$\epsilon_t = X_t - \mu_t$$

is the unexpected return at time t.

In GARCH–Jump model it is supposed that news process have two separate components (normal and unusual news), which cause two types of innovation (smooth and jump-like innovations):

$$\epsilon_t = \epsilon_{1,t} + \epsilon_{2,t}. \tag{7.1}$$

These two news innovations have a different impact on return volatility. It is assumed that the first component $\epsilon_{1,t}$ reflects the impact of unobservable normal news innovations, while the second one $\epsilon_{2,t}$ is caused by unusual news events.

The first term in (7.1) reflects the impact of normal news to volatility:

$$\epsilon_{1,t} = \sigma_t u_t, t \in \mathbb{Z},$$

where (u_n) be a sequence of i.i.d. random variables such that $u_t \sim \mathcal{N}(0,1)$, (σ_t) is a nonnegative GARCH(1,1) process such that

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

and $\alpha_0, \alpha_1, \beta_1 > 0$. Note that $\mathbb{E}(\epsilon_{1,t}|I_{t-1}) = 0$.

The second term in (7.1) is a jump innovation with $\mathbb{E}(\epsilon_{2,t}|I_{t-1}) = 0$. The component $\epsilon_{2,t}$ is a result of unexpected events and is responsible for jumps in volatility.

The distribution of jumps is assumed to be Poisson distribution. Let λ_t be intensity parameter of Poisson distribution. Denote n_t a number of jumps occurring between time $t-1$ and t. Then conditional density of n_t is

$$P(n_t = j|I_{t-1}) = \frac{\exp(-\lambda_t)\lambda_t^j}{j!}, \ j = 0, 1, \ldots. \tag{7.2}$$

We suppose that the intensity parameter λ_t conditionally varies over time. It is assumed that the conditional jump intensity $\lambda_t = \mathbb{E}(n_t|I_{t-1})$, i.e. the expected number of jumps occurring between time $t-1$ and t conditional on information I_{t-1}, has dynamics

$$\lambda_t = \lambda_0 + \rho\lambda_{t-1} + \gamma_1\zeta_{t-1}. \tag{7.3}$$

The process (7.3) is called an autoregressive conditional jump intensity and was proposed in the paper Chan and Maheu [2002]. The model based on the assumption that the conditional jump intensity is autoregressive and related both to the last period's conditional jump intensity and to an intensity residual ζ_{t-1}. The intensity residual ζ_{t-1} is defined as

$$\zeta_{t-1} = \mathbb{E}(n_{t-1}|I_{t-1}) - \lambda_{t-1} = \sum_{j=0}^{\infty} jP(n_{t-1} = j|I_{t-1}) - \lambda_{t-1}.$$

Here $\mathbb{E}(n_{t-1}|I_{t-1})$ is the expected number of jumps occurring from $t-2$ to $t-1$, and λ_{t-1} is the conditional expectation of numbers of jumps n_{t-1} given the information I_{t-2} available at the moment $t-2$. Thus

$$\zeta_{t-1} = \mathbb{E}(n_{t-1}|I_{t-1}) - \mathbb{E}(n_{t-1}|I_{t-2})$$

i.e. ζ_{t-1} represents the change in the econometrician's conditional forecast of n_{t-1} as the information set is updated from $t-2$ to $t-1$. It is easy to see that $\mathbb{E}(\zeta_t|I_{t-1}) = 0$, i.e. ζ_t is a martingale difference sequence with respect to I_{t-1}, and therefore $\mathbb{E}(\zeta_t) = 0$, $\mathrm{Cov}(\zeta_t, \zeta_{t-i}) = 0$ for all $i > 0$.

Denote $Y_{t,k}$ the size of k-th jump that occur from time $t-1$ to t, $1 \le k \le n_t$. In the model it is supposed that the jump size $Y_{t,k}$ is realization of normal distributed random:

$$Y_{t,k} \sim \mathcal{N}(\theta, \delta^2).$$

Then the cumulative jump size J_t from $t-1$ to t is equal to the sum of all jumps occurring from time $t-1$ to t:

$$J_t = \sum_{k=1}^{n_t} Y_{t,k}.$$

The jump innovation $\epsilon_{2,t}$ defined by

$$\epsilon_{2,t} = J_t - \mathbb{E}(J_t|I_{t-1}).$$

It follows from

$$\mathbb{E}(J_t|I_{t-1}) = \theta\lambda_t$$

that

$$\epsilon_{2,t} = \sum_{k=1}^{n_t} Y_{t,k} - \theta\lambda_t.$$

Therefore we have

$$\mathbb{E}(\epsilon_{2,t}|I_{t-1}) = 0.$$

7.1.2 Maximum Likelihood Estimation of GARCH Model with Jumps

The subsection describes quasi-maximum likelihood estimation (QML) of GARCH model with Jumps. The vector of model parameters is

$$\Theta = (\alpha_0, \alpha_1, \beta_1, \delta, \theta, a, b, c)^T.$$

We will assume that θ belongs to the set

$$S := \{(\alpha_0, \alpha_1, \beta_1, \delta, \theta, a, b, c)^T : \ \alpha_0 \geq 0, \ \alpha_1 > 0, \ \beta_1 > 0\}.$$

Denote

$$\Theta^* = (\alpha_0^*, \alpha_1^*, \beta_1^*, \delta^*, \theta^*, a^*, b^*, c^*)^T$$

the vector of the true values of parameters. The aim is to find Θ^* that maximize a QML function given an observation sequence

$$\epsilon_0, \ldots, \epsilon_n$$

of length n.

Define the sequence $(\tilde{\sigma}_1, \ldots, \tilde{\sigma}_n)$ by recursion:

$$\tilde{\sigma}_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \tilde{\sigma}_{t-1}^2.$$

If we assume that the likelihood function is Gaussian, then the log-likelihood function can be written as (see e.g. Chan and Maheu [2002]):

$$F_n(\Theta) := \sum_{t=1}^{n} \log f(\epsilon_t | I_{t-1}, \Theta),$$

where

$$f(\epsilon_t | I_{t-1}, \Theta) = \sum_{j=0}^{\infty} \frac{\exp(-\tilde{\lambda}_t)\tilde{\lambda}_t^j}{j!} f(\epsilon_t | n_t = j, I_{t-1}, \Theta) \tag{7.4}$$

and

$$f(\epsilon_t | n_t = j, I_{t-1}, \Theta) = \frac{1}{\sqrt{2\pi(\tilde{\sigma}_t^2 + j\delta^2)}} \exp\left(-\frac{(\epsilon_t + \theta\lambda_t - \theta j)^2}{2(\tilde{\sigma}_t^2 + j\delta^2)} \right). \tag{7.5}$$

The sequence of $\tilde{\lambda}_t$ is defined by recursion:

$$\tilde{\lambda}_t = a + b\tilde{\lambda}_{t-1} + c\zeta_{t-1},$$

where

$$\zeta_{t-1} = \mathbb{E}(n_{t-1}|I_{t-1}) - \tilde{\lambda}_{t-1},$$

and

$$\begin{aligned}
\mathbb{E}(n_{t-1}|I_{t-1}) &= \sum_{j=0}^{\infty} jP(n_{t-1} = j|I_{t-1}) = \\
&= \sum_{j=0}^{\infty} j\frac{f(\epsilon_t|n_{t-1} = j, I_{t-2}, \Theta)P(n_{t-1} = j)|I_{t-2})}{f(\epsilon_t|I_{t-2}, \Theta)} = \\
&= \frac{\sum_{j=1}^{\infty} \frac{\exp(-\tilde{\lambda}_{t-1})\tilde{\lambda}_{t-1}^j}{j!} \frac{1}{\sqrt{2\pi(\tilde{\sigma}_{t-1}^2 + j\delta^2)}} \exp\left(-\frac{(\epsilon_{t-1}+\theta\lambda_{t-1}-\theta j)^2}{2(\tilde{\sigma}_{t-1}^2+j\delta^2)}\right)}{f(\epsilon_{t-1}|I_{t-2}, \Theta)}
\end{aligned} \tag{7.6}$$

The maximum likelihood estimator of Θ is defined by

$$\Theta^* = \arg\max_{\Theta \in S} F_n(\Theta).$$

Since the densities (7.5) has an infinite sum, it is impossible to use them for parameters' estimation. There are two ways of using equation (7.5):

- taking a finite Taylor expansions of (7.5);
- truncation of the sum (7.5), i.e. limitation of the number of terms in the sum.

MATLAB code for calibration the GARCH model with jumps is in Appendix B. It is should be noted that the calibration problem is non convex and surface of optimized function has a highly complex relief. As it was mentioned in Chapter 3, finding its exact solution is a difficult task. We faced with difficulties when calibrate process via MATLAB function `fminsearch`. In particular, the calibration process is not robust and extremely sensitive to the choice of a starting point. For this reason, we do not include any empirical results for the GARCH model with jumps (the case of autoregressive jump intensity). However, if we would assume that jump intensity is constant over time then the calibration process converges. MATLAB code for calibration of the GARCH model with constant jump intensity also can be found in Appendix B.

7.1.3 Empirical Results

We use the data set described in Section 6.3. Dataset includes the daily stock closing prices of five companies traded on London Stock Exchange: AVIVA, BP, BT Group, Lloyd Banking Group, HSBC.

Table 6.2 shows preliminary descriptive statistics for the stock prices log returns.

Table 7.1 shows the maximum likelihood estimates of GARCH(1,1) model with Jumps (with constant jump intensity, i.e. it is assumed that $b = c = 0$) for log returns of the closing daily prices of the five companies for 3 years (July 5, 2005 - July 5, 2008).

Table 7.1: Maximum likelihood estimates of GARCH(1,1) model with Jumps for log returns of the closing daily prices

Company	α	β	δ	θ	λ	$\alpha + \beta$	LLF_6
AVIVA	0.1247	0.8248	1.44E-02	-9.66E-03	0.9496	0.9495	2804.88
BP	0.0918	0.7919	1.02E-02	4.95E-04	0.8837	0.8837	2875.06
BT Group	0.0406	0.9332	1.87E-02	1.05E-03	0.9738	0.9738	2825.57
Lloyds	0.1262	0.8464	1.45E-02	4.11E-04	0.9726	0.9726	2899.96
HSBC	0.1335	0.8278	1.56E-02	-6.52E-04	0.9613	0.9613	3126.34

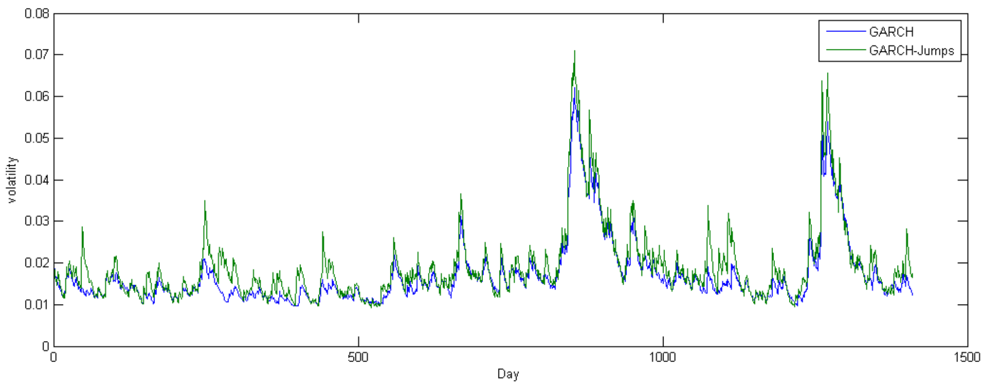

Figure 7.1: GARCH model and GARCH model with Jumps performance for BP stock market closing daily prices (January 5, 2005 - December 31, 2010)

7.2 Individual Stock Volatility Modelling With GARCH–Jumps Model Augmented With News Analytics Data

7.2.1 Model description

We are going to analyze the impact of news process intensity on stock volatility by extending GARCH-Jump models proposed and studied in Maheu and McCurdy [2004].

Let X_t be the log return of a particular stock or the market portfolio from time $t-1$ to time t. Let I_{t-1} denotes the past information set containing the realized values of all relevant variables up to time $t-1$. Suppose investors know the information in I_{t-1} when they make their investment decision at time $t-1$. Then the relevant expected return μ_t to the investors is the conditional expected value of X_t, given I_{t-1}, i.e.

$$\mu_t = E(X_t | I_{t-1}).$$

The relevant expected volatility σ_t^2 to the investors is conditional variance of X_t, given I_{t-1}, i.e.

$$\sigma_t^2 = Var(X_t | I_{t-1}).$$

Then

$$\epsilon_t = X_t - \mu_t$$

is the unexpected return at time t. Following Maheu and McCurdy [2004] we suppose that news process have two separate components: normal and unusual news,

$$\epsilon_t = \epsilon_{1,t} + \epsilon_{2,t}. \tag{7.7}$$

The first term in (7.7) reflects the impact of normal news to volatility:

$$\epsilon_{1,t} = \sigma_t u_t, t \in \mathbb{Z},$$

where (u_n) be a sequence of i.i.d. random variables such that $u_t \sim N(0,1)$, (σ_t) is a nonnegative process such that

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

and

$$\alpha_0, \alpha_1, \beta_1 > 0.$$

The second term in (7.7) reflects the result of unexpected events and describe jumps in volatility:

$$\epsilon_{2,t} = \sum_{k=1}^{N_t} Y_{t,k} - \theta\lambda_t,$$

where $Y_{t,k} \sim \mathcal{N}(\theta, \delta^2)$, N_t is a Poisson random variable with conditional jump intensity

$$\lambda_t = a + b\lambda_{t-1} + c\zeta_{t-1} + \rho_1 n^+_{t-1} + \rho_2 n^-_{t-1},$$

where $\zeta_{t-1} = \mathbb{E}(N_{t-1}|I_{t-1}) - \theta\lambda_{t-1}$, and n^+_{t-1}, n^-_{t-1} is the number of positive and negative news from $t-2$ to $t-1$ respectively. Therefore we directly take into account the qualitative data of news intensity and news sentiment score (source: RavenPack News Scores).

7.2.2 Empirical results

Table 7.2 presents maximum likelihood estimates of GARCH(1,1)–Jumps model augmented with news intensity for log returns of the closing daily prices for the five companies (January 5, 2005 - December 31, 2010). It shows that $\rho_1 < \rho_2$ for all companies, i.e. the impact of the number of negative news on the growth of jump intensity much higher than one's of positive news.

Table 7.2: Maximum likelihood estimates of GARCH(1,1)–Jumps model augmented with news intensity for log returns of the closing daily prices

Company	α	β	δ	θ	λ	ρ_1	ρ_2	LLF_7
AVIVA	0.12	0.82	1.4E-02	-9.7E-03	0.14	0.011	0.12	2876.37
BP	0.09	0.79	1.0E-02	4.9E-04	0.58	0.032	0.42	3239.31
BT Group	0.04	0.93	1.9E-02	1.0E-03	0.26	0.03	0.42	2835.06
Lloyds	0.13	0.85	1.4E-02	4.1E-04	0.20	0.04	0.13	2909.35
HSBC	0.13	0.83	1.6E-02	-6.5E-04	0.06	0.00	0.01	3128.33

Note that the GARCH model with jumps (the null model) is a special case of the augmented GARCH-Jumps model (the alternative model). Therefore, to compare the fit of two models it can be used a likelihood ratio test (see e.g. Cox and Hinkley [1974]). Results of likelihood ratio test are in Table 7.3. For tree of five companies the alternative model is preferable with confidence level 5%.

Table 7.3: Results of the likelihood ratio test for the GARCH model with jumps and the augmented GARCH-Jumps model

Company	LLF_6	LLF_7	$2(LLF_7 - LLF_6)$	$\chi^2(2), 5\%$	Null Hyp.
AVIVA	2804.89	2876.37	142.96	5.99	rejected
BP	2875.06	3239.31	728.50	5.99	rejected
BT Group	2825.58	2835.06	18.96	5.99	rejected
Lloyds	2899.97	2909.35	18.77	5.99	rejected
HSBC	3126.34	3128.33	3.98	5.99	accepted

7.3 Summary

In the chapter we have examined two GARCH models with jumps. First we consider the well-known GARCH model with jumps proposed in Maheu and McCurdy [2004]. Then we introduced the GARCH–Jumps model augmented with news intensity and obtained some empirical results. The main assumption of the model is that jump intensity might change over time and that jump intensity depends linearly on the number of positive and negative news. It is not clear whether news adds any value to a jump-GARCH model. However, the comparison of the values of log likelihood shows that the GARCH–Jumps model augmented with news intensity performs slightly better than ″pure″ GARCH or the GARCH model with Jumps.

Chapter 8

Summary and future work

8.1 Summary and contributions

The first part of the thesis compares the performance of different GARCH models using backtesting procedures based on unconditional coverage test (Kupiec's test) and the test of conditional coverage.

The data set we have analyzed in this part of the work is the stock market closing daily prices of General Electric Company (GEC.L). The sample period is from January 2, 2008 to December 31, 2010. Data set are taken from UK Stock Market FTSE100 and downloaded from Yahoo!Finance site. The sample is divided in twelve parts for two purposes: in-sample estimation procedure and out-of-sample evaluation. Using the Jarque–Bera test statistics with 1% level we reject the null hypothesis of normality of log return series. Shapiro–Wilk W test also rejects the hypothesis of normal distribution of the data.

It was shown that the ARCH(1), GARCH (1, 1), TGARCH(1,1) and GJR-GARCH (1, 1) models calibrated on data sets of 1-year length (January 2, 2009 - December 31, 2009) under the normal distribution performed well. However, for 95% and 99% VaR estimations all the ARCH(1), GARCH (1, 1), TGARCH(1,1) and GJR-GARCH (1, 1) models calibrated on datasets of 2-year length (January 2, 2008 - December 31, 2009) underestimated the risk and was rejected. It might be explained by huge difference in the level of volatility in 2008 crisis year compare with the one in 2009 year.

The second part of the work tries to evaluate the impact of news on stock volatility. There are not so much research works studying quantitative impact of news on stock volatility. It is worth to be mentioned the pioneering works Kalev et al. [2004] and Janssen [2004]. In the second of the papers the author examines impact of news releases on *index* volatility, while in our work we analyze the impact on *stock* volatility following study of Kalev et al. [2004]. However, we restrict our choice by some of the FTSE100 companies, while Kalev et al. [2004] considered some French companies.

In the Chapter 5.5 we have tried to study different GARCH models augmented with news analytics data. The main goal was to examine the impact of news intensity on stock volatility. Based on empirical evidences for some of FTSE100 companies it has been shown shown that the GARCH(1,1) model augmented with volume does remove GARCH and ARCH effects for most of the FTSE100 companies, while the GARCH(1,1) model augmented with news intensity has difficulties in removing the effects. It has been shown that the GARCH(1,1) model augmented with the news intensity n_t (the number of daily announcements) does not necessarily remove GARCH and ARCH effects. However, the likelihood ratio test has shown that the GARCH(1,1) model augmented with the news intensity performs better than the "pure" GARCH model.

This study uses Monte Carlo hypothesis tests with the log likelihood ratio as the test statistic. We use Monte Carlo methods to obtain the probability of a larger value of the test statistic under the null hypothesis. Based on the maximum likelihood estimation technique, we estimate two competing time series models (GARCH model and augmented GARCH models) of daily prices of five FTSE100 company. Using Monte Carlo hypothesis tests, we conclude that

- for 4 of 5 companies the GARCH-volume model cannot be rejected, while the GARCH model is rejected;
- for all companies, the GARCH-lagged-volume model must be rejected;
- for 4 of 5 companies, the GARCH-news model cannot be rejected, while the GARCH model is rejected;

- for 3 of 5 companies, the GARCH-lagged-news model cannot be rejected, while the GARCH model is rejected.

Then we compare GARCH model with jumps and GARCH–Jumps model augmented with news intensity using likelihood ratio test.

To calibrate the models we have used the Maximum Likelihood Estimation (MLE) and Quasi Maximum Likelihood Estimation (QMLE) methods. The volatility models are calibrated on the software package MATLAB. We used RavenPack news analytics data.

8.2 Future work

The study has shown that the problem of examining the impact of news intensity on stock volatility is far more sophisticated than it might seem at first sight. The empirical results show that there are no strong arguments in support of hypothesis of impact of news intensity on volatility. It might be occur due the following causes:

- an additional preprocessing of news analytics data is required;
- the models do not take into account macro economics news;
- it could be seen the splash of volatility in 2008, while news intensity has not been changing very much over the whole period;
- news intensity might not affect volatility directly, although jumps in returns could be caused by news releases.

The work may be considered as a preliminary work on the problem of evaluation of impact of news on stock volatility. Based on the research it can be suggested some directions of future work.

- The first problem is to develop a GARCH-type model with news analytics data for prediction VaR with better performance than the "pure" GARCH model.

- It is worth considering the problem of mutual dependence of volatility and news intensity.
- The problem of calibration of augmented models (e.g. GARCH–Jumps models) is difficult due to its non convexity and noisiness (the problem was mentioned in Chapter 7). We can try to use different solvers for global optimization or to develop new algorithms.

Future work may be also associated with the study of

- *Markov – Switching GARCH models.* The idea is to estimate a model that permits regime switching in the parameters caused by movements of news intensity. It is a generalization of the GARCH model and permits a different persistence in the conditional variance of each regime. Thus, the conditional variance in each regime accommodates volatility clustering, nesting the GARCH model as special case.
- *HMM – GARCH Model*. The model is similar to the previous one, but it is supposed that the process is a hidden Markov process. We will suppose that the hidden states in HMM are "somehow" connected with observable sequence of the news sentiment score and parameters of GARCH model are state-dependant.

There are some evidences (see e.g. Mitra and Mitra [2011]) that effect of news on prices is short-term, therefore it is more likely that we need tick by tick data to examine impact of news on stock volatility.

Appendix A

Tables

ID dataset	From	To	Size
0	02/01/2008	31/12/2010	577
1yy	02/01/2008	31/12/2009	505
2yy	01/02/2008	02/02/2010	505
3yy	04/03/2008	04/03/2010	505
4yy	03/04/2008	05/04/2010	505
5yy	01/05/2008	04/05/2010	505
6yy	03/06/2008	03/06/2010	505
7yy	02/07/2008	02/07/2010	505
8yy	01/08/2008	03/08/2010	505
9yy	02/09/2008	01/09/2010	505
10yy	01/10/2008	01/10/2010	505
11yy	30/10/2008	01/11/2010	505
12yy	28/11/2008	01/12/2010	505
1y	02/01/2009	31/12/2009	252
2y	01/02/2009	02/02/2010	252
3y	04/03/2009	04/03/2010	252
4y	03/04/2009	05/04/2010	252
5y	01/05/2009	04/05/2010	252
6y	03/06/2009	03/06/2010	252
7y	02/07/2009	02/07/2010	252
8y	01/08/2009	03/08/2010	252
9y	02/09/2009	01/09/2010	252
10y	01/10/2009	01/10/2010	252
11y	30/10/2009	01/11/2010	252
12y	28/11/2009	01/12/2010	252

Table A.1: In-sample data sets

ID dataset	From	To	Size
A	04/01/2010	02/02/2010	21
B	03/02/2010	04/03/2010	21
C	05/03/2010	05/04/2010	21
D	06/04/2010	04/05/2010	21
E	05/05/2010	03/06/2010	21
F	04/06/2010	02/07/2010	21
G	06/07/2010	03/08/2010	21
H	04/08/2010	01/09/2010	21
K	02/09/2010	01/10/2010	21
L	04/10/2010	01/11/2010	21
M	02/11/2010	01/12/2010	21
N	02/12/2010	31/12/2010	21

Table A.2: Out-of-sample data sets

ID dataset	Size	Mean	Std.dev.	Min	Max	Skew	Kurt
0	756	-.092	3.05	-13.68	17.98	.05	7.71
1yy	504	-.177	3.53	-13.68	17.98	.11	6.33
2yy	505	-.148	3.53	-13.68	17.98	.099	6.34
3yy	505	-.144	3.53	-13.68	17.98	.096	6.33
4yy	505	-.142	3.51	-13.68	17.98	.086	6.46
5yy	505	-.114	3.47	-11.81	17.98	.174	6.29
6yy	505	-.122	3.51	-11.81	17.98	.177	6.05
7yy	505	-.133	3.51	-11.81	17.98	.188	6.05
8yy	505	-.108	3.56	-11.81	17.98	.169	6.09
9yy	505	-.124	3.50	-11.81	17.98	.179	6.12
10yy	505	-.088	3.37	-11.81	17.98	.175	6.70
11yy	505	-.037	3.19	-11.81	17.98	.147	7.27
12yy	505	-.010	2.96	-11.38	17.98	.333	8.21
1y	253	-.019	3.55	-11.38	17.98	.362	6.56
2y	253	.143	3.32	-11.31	17.98	.576	7.61
3y	253	.353	2.86	-8.76	17.98	1.05	9.44
4y	253	.217	2.29	-8.76	6.82	-.094	4.31
5y	253	.128	2.11	-6.23	6.70	.09	3.81
6y	253	.059	2.13	-6.24	6.70	.165	4.03
7y	253	.078	2.10	-6.23	6.70	.230	4.05
8y	253	.072	1.97	-5.96	6.64	.151	3.93
9y	253	.050	1.93	-5.97	6.64	.258	4.083
10y	253	.003	1.82	-5.97	6.64	.101	4.36
11y	253	.044	1.75	-5.97	6.64	.102	4.81
12y	253	.0123	1.72	-5.97	6.65	-.0151	4.65

Table A.3: Descriptive statistics of log returns of GE stock market closing daily prices

LAG	Auto. corr	PAC	Q	$Prob > Q$
1	-0.016	-0.016	0.19481	0.6589
2	0.023	0.0228	0.59748	0.7418
3	0.0066	0.0074	0.63086	0.8893
4	0.0414	0.0412	1.9398	0.7468
5	-0.0683	-0.0675	5.4942	0.3586
6	0.0676	0.0642	8.9892	0.1742
7	0.0126	0.0167	9.1113	0.2448
8	-0.0441	-0.0478	10.601	0.2253
9	-0.0595	-0.0572	13.32	0.1486
10	0.0874	0.0795	19.19	0.0379

Table A.4: The autocorrelations, partial autocorrelations and Portmanteau (Q) statistics for GE log returns of the closing daily prices

ID dataset	α_1	α_0	LLF
0	.45318247* (.0501394)	.00056253* (.0000239)	1623.207
1yy	.43468087* (.0648449)	.0007765* (.0000455)	1001.271
2yy	.43616612* (.0651583)	.00077712* (.0000451)	1003.371
3yy	.42863965* (.064351)	.00078176* (.0000454)	1003
4yy	.45254544* (.0658052)	.00075293* (.0000432)	1008.561
5yy	.47683755* (.0669086)	.0007063* (.0000434)	1018.169
6yy	.42417881* (.0637027)	.00076458* (.000049)	1006.906
7yy	.41995754* (.0632787)	.00076729* (.0000492)	1006.732
8yy	.42054099* (.0632339)	.00076523* (.0000488)	1007.755
9yy	.43315058 (.0641537)	.00075466 (.0000478)	1009.334
10yy	.45160768* (.0623002)	.00068607* (.0000412)	1031.405
11yy	.51300839* (.0639891)	.0005694* (.0000341)	1066.638
12yy	.54630221* (.0649232)	.00049519* (.000033)	1099.165

Table A.5: ARCH(1) model estimates of returns for the GE stock market closing daily prices, 2-year length data sets, (January 2, 2008 - December 31, 2010), $\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2$.
Standard errors appear in parentheses.
* Statistically significant at 5% assuming that returns are conditionally normally distributed.

ID dataset	α_1	α_0	LLF
1y	0.47094295* (.0897237)	0.00077797* (.000082)	504.4978
2y	0.46953273* (.0825078)	0.0006858* (.0000643)	521.5608
3y	.5176217* (.0765507)	.00052667* (.0000478)	551.1712
4y	.10449969 (.0727847)	.00050226* (.0000412)	596.7184
5y	.11002752 (.0834146)	.00039972* (.0000394)	625.2793
6y	.09713145 (.083599)	.00040447* (.000038)	625.1935
7y	.07077441 (.0770472)	.00040942* (.0000378)	626.681
8y	.13831818 (.085816)	.00035463* (.0000343)	637.3311
9y	.14835698 (.0802731)	.00032495* (.0000297)	647.6595
10y	.0690674 (.0744808)	.00030922* (.0000273)	662.8278
11y	.05540686 (.0500281)	.00030167* (.0000232)	667.6224
12y	.00300171 (.0498605)	.00029302* (.0000234)	679.7484

Table A.6: ARCH(1) model estimates of returns for the GE stock market closing daily prices, 1-year length data sets (January 2, 2008 - December 31, 2010), $\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2$
Standard errors appear in parentheses.
* Statistically significant at 5% assuming that returns are conditionally normally distributed.

ID dataset	ω	α	β	$\alpha + \beta$	LLF
0	.0000117* (4.11e-06)	0.091613* (.0166542)	0.89478462* (.017231)	0.98639762	1718.774
1yy	.0000208* (9.13e-06)	0.11774579 (.0258138)	0.86870348 (.0270232)	0.98644927	1052.985
2yy	.0000261* (1.00e-05)	0.1123373* (.0257945)	0.86790682* (.0282388)	0.98024412	1053.44
3yy	.0000108* (4.50e-06)	0.07642264* (.0120821)	0.91316292* (.0115348)	0.98958556	1057.965
4yy	.0000126* (4.26e-06)	0.10249227* (.0149669)	0.88411541* (.0141541)	0.98660768	1074.261
5yy	.0000167 (6.04e-06)	0.15243927* (.0365819)	0.84018656* (.0334933)	0.99262583	1091.03
6yy	.0000142* (5.65e-06)	0.12991837* (.0292203)	0.86236451* (.0281027)	0.99228288	1074.575
7yy	.0000146* (5.70e-06)	0.13769394* (.0329878)	0.85413146* (.0314031)	0.9918254	1075.705
8yy	.0000144* (5.75e-06)	0.13433309* (.0325279)	0.85729065* (.0311362)	0.99162374	1077.227
9yy	9.75e-06* (4.67e-06)	0.09679123* (.0259928)	0.8932116* (.0268273)	0.99000283	1082.436
10yy	8.20e-06* (4.04e-06)	0.08806305* (.0225051)	0.90025755* (.0241458)	0.9883206	1107.475
11yy	9.60e-06* (4.23e-06)	0.0896181* (.0225954)	0.89576975* (.0249585)	0.98538785	1134.459
12yy	8.82e-06* (4.00e-06)	0.08074322* (.0195418)	0.90290047* (.0231342)	0.98364369	1163.586

Table A.7: GARCH(1,1) model estimation of returns of the GE company closing daily prices, 2-year length of datasets, (January 2, 2008 - December 31, 2010), $\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$
Standard errors appear in parentheses.
* Statistically significant at 5% assuming that returns are conditionally normally distributed.

ID dataset	ω	α	β	$\alpha+\beta$	LLF
1y	-0.00049865 (6.53e-06)	0.11066711* (.0398877)	0.89575842* (.0380215)	1.00642553	536.262
2y	0.00150529 (9.54e-06)	0.09006688* (.0398533)	0.89169194* (.0465566)	0.98175882	550.0163
3y	0.00069579 (.0000106)	0.12899038* (.0401374)	0.84032848* (.0490059)	0.96931886	575.7954
4y	0.00231659 (1.79e-06)	0.01650656* (.0074601)	0.98261196* (.0099026)	0.99911852	612.3225
5y	0.00147233 (.0000121)	0.09654812* (.043362)	0.86018358* (.0598358)	0.95673170	633.1722
6y	0.00106784 (.0000127)	0.10907783* (.0437474)	0.84420851* (.0592344)	0.95328634	633.1432
7y	0.00077362 (.0000126)	0.10472511* (.0427512)	0.84419683* (.0595142)	0.94892194	634.908
8y	0.00136975 (.0000147)	0.08322212* (.0382842)	0.85573793* (.0653853)	0.93896005	641.8829
9y	0.00050821 (.0000174)	0.09019922* (.0412902)	0.83197183* (.0750523)	0.92217105	652.0249
10y	0.00069545 (.000016)	0.07540999* (.036644)	0.84374419* (.0744698)	0.91915418	667.5812
11y	0.00070645 (.0000167)	0.0703* (.0345722)	0.83669513* (.0783554)	0.90699513	672.6497
12y	0.00054304 (.0000144)	0.06192434* (.0295961)	0.86216198* (.0701283)	0.92408632	685.4497

Table A.8: GARCH(1,1) model estimation of returns of the GE company closing daily prices, 1-year length of datasets, (January 2, 2008 - December 31, 2010), $\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$
Standard errors appear in parentheses.
* Statistically significant at 5% assuming that returns are conditionally normally distributed.

Dataset ID	α_0	α^+	α^-	β_1	LLF
0	0.000554 (.0007157)	0.047521* (.0125592)	0.135362* (.0244669)	0.908928* (.0182362)	1732.939
1y	.0000208 (.0014961)	0.048771* (.0148612)	0.138605* (.0471608)	0.930055* (.0280554)	541.7385
2y	0.000200 (.0015109)	0.017903 (.0106075)	0.107432* (.0357879)	0.942072* (.0255355)	555.6529
3y	0.000377 (.0013158)	0.074240* (.0204521)	0.158485* (.0413718)	0.895850* (.033684)	577.8982
4y	-0.000026 (.0013032)	0.027466* (.0029056)	0.038922* (.0306135)	0.972819* (.0178327)	612.0834
5y	0.001379 (.0011526)	0.117877* (.027686)	0.136405* (.0624604)	0.837491* (.0677835)	634.7368
6y	0.001002 (.001122)	0.067946* (.0220407)	0.137234* (.053377)	0.875547* (.0528859)	636.0632
7y	0.001294 (.0011547)	0.096650* (.0385069)	0.144320* (.059806)	0.846322* (.0610073)	637.2429
8y	0.001303 (.0011155)	0.066769* (.0136125)	0.125496* (.0487081)	0.859332* (.0547106)	644.9564
9y	0.001346 (.0010978)	0.062836* (.021556)	0.120286* (.0469466)	0.859621* (.0563033)	654.6672
10y	0.001456 (.0010509)	-0.062297* (.0176594)	0.090683* (.0253895)	0.906351* (.0347344)	673.0868
11y	0.001524 (.0010059)	0.002831 (.0086602)	0.111050* (.0450289)	0.868877* (.0531347)	676.6295
12y	0.001402 (.0010174)	0.020156* (.0079958)	0.102713* (.0369986)	0.870911* (.0535984)	688.3637

Table A.9: Threshold GARCH(1,1) model estimation of returns of the GE company closing daily prices (January 2, 2008 - December 31, 2010), $\sigma_t = \alpha_0 + \alpha^+\epsilon^+_{t-1} + \alpha^-\epsilon^-_{t-1} + \beta_1\sigma_{t-1}$
Standard errors appear in parentheses.
* Statistically significant at 5% assuming that returns are conditionally normally distributed.

DatasetID	α_0	α_1	β_1	γ_1	LLF
0	0.0001 (.0008329)	0.1190* (.0242256)	0.8902* (.0222389)	-0.0592* (.0225011)	1720.4640
1yy	-0.0019 (.0012853)	0.1479* (.0379002)	0.8632* (.0331379)	-0.0648 (.0389331)	1053.8720
2yy	0.0013 (.0012947)	0.1530* (.0426398)	0.8616* (.0357976)	-0.0878* (.0421753)	1055.0530
3yy	0.0011 (.0012317)	0.0887* (.0128993)	0.9282* (.0124403)	-0.0609* (.0199828)	1059.9700
4yy	0.0005 (.0011372)	0.1330* (.0236991)	0.9027* (.0139189)	-0.1068* (.0300345)	1078.6960
5yy	0.0005 (.0010362)	0.1675* (.0443843)	0.8946* (.0274805)	-0.1504* (.0424582)	1097.8390
6yy	0.0003 (.0010526)	0.1499* (.0363991)	0.9115* (.0233486)	-0.1498* (.0358784)	1083.4730
7yy	0.0005 (.0010551)	0.1554* (.0393737)	0.9055* (.0257752)	-0.1506* (.0376036)	1084.2460
8yy	0.0001 (.0010413)	0.1504* (.0387997)	0.9086* (.024837)	-0.1479* (.0374164)	1086.2130
9yy	-0.0003 (.0010046)	0.1172* (.0308649)	0.9284* (.020328)	-0.1208* (.0310307)	1091.4110
10yy	0.0001 (.0009482)	0.1136* (.0292088)	0.9277* (.0198916)	-0.1154* (.0292727)	1116.6760
11yy	0.0001 (.0009273)	0.1055* (.0271074)	0.9260* (.0202897)	-0.0992* (.0277119)	1140.5850
12yy	0.0002 (.0008879)	0.1044* (.0251986)	0.9255* (.0196989)	-0.0959* (.026294)	1169.5120

Table A.10: GJR–GARCH(1,1) model estimation of returns of GE company closing daily prices (January 2, 2008 - December 31, 2010), 2-year lengh of datasets, $\sigma_t^2 = \alpha_0 + \alpha_1\epsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2 + \gamma_1\epsilon_{t-1}^2\chi_{t-1}$, where $\chi_{t-1} = 0$ if $\epsilon_{t-1} \geq 0$, and $\chi_{t-1} = 1$ if $\epsilon_{t-1} < 0$
Standard errors appear in parentheses.
* Statistically significant at 5% assuming that returns are conditionally normally distributed.

DatasetID	α_0	α_1	β_1	γ_1	LLF
1y	-0.0015 (.0014732)	0.1576* (.0639027)	0.9160* (.0346217)	-0.1271* (.0558179)	541.3910
2y	0.0007 (.0015879)	0.1220* (.0524979)	0.9217* (.0367099)	-0.1123* (.0490246)	553.7450
3y	0.0002 (.001435)	0.1433* (.0479758)	0.8947* (.0378663)	-0.1076* (.0470933)	577.7362
4y	0.0017 (.0013037)	0.0335* (.0167137)	0.9872* (.0083994)	-0.0450 (.0295633)	613.2887
5y	0.0015 (.0012378)	0.0623* (.0360488)	0.9434* (.0345582)	-0.0424 (.0377764)	633.7611
6y	0.0006 (.0012845)	0.0987* (.0306063)	0.9505* (.0257079)	-0.1140* (.0350878)	636.2532
7y	0.0006 (.0012032)	0.1305* (.0627152)	0.8814* (.0560835)	-0.0913 (.0585258)	636.2303
8y	0.0011 (.0011415)	0.1109* (.0531542)	0.8879* (.0543036)	-0.0863 (.0505297)	643.6099
9y	0.0004 (.0011445)	0.1147* (.0566202)	0.8756* (.0599397)	-0.0882 (.0532909)	653.5250
10y	0.0002 (.0010541)	0.0918* (.0210626)	0.9321* (.025575)	-0.1766* (.0295376)	675.8738
11y	0.0005 (.0010187)	0.0908* (.0266322)	0.9211* (.0352009)	-0.1530* (.0324703)	677.6904
12y	0.0004 (.0010443)	0.0911* (.0332124)	0.8951* (.0446269)	-0.1346* (.0482576)	687.9426

Table A.11: GJR–GARCH(1,1) model estimation of returns of GE company closing daily prices (January 2, 2008 - December 31, 2010), 1-year length of datasets, $\sigma_t^2 = \alpha_0 + \alpha_1\epsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2 + \gamma_1\epsilon_{t-1}^2\chi_{t-1}$, where $\chi_{t-1} = 0$ if $\epsilon_{t-1} \geq 0$, and $\chi_{t-1} = 1$ if $\epsilon_{t-1} < 0$
Standard errors appear in parentheses.
* Statistically significant at 5% assuming that returns are conditionally normally distributed.

Model	Length	Conf	Failures	n_{00}	n_{01}	n_{10}	n_{11}	Π_0	Π_1	Π
ARCH(1,1)	1 year	95%	8	235	7	7	1	0.032089	0.1250	0.0320
GARCH(1,1)	1 year	95%	9	232	9	9	0	0.0373	0.0000	0.0360
GJR-GARCH	1 year	95%	10	232	10	10	0	0.0413	0.0000	0.0397
TGARCH	1 year	99%	10	232	10	10	0	0.0413	0.0000	0.0397
ARCH(1,1)	1 year	99%	2	246	2	2	0	0.0081	0.0000	0.0080
GARCH(1,1)	1 year	99%	3	244	3	3	0	0.0121	0.0000	0.0120
GJR-GARCH	1 year	99%	3	244	3	3	0	0.0121	0.0000	0.0120
TGARCH	1 year	99%	3	244	3	3	0	0.0121	0.0000	0.0120
ARCH(1,1)	2 year	95%	1	248	1	1	0	0.0040	0.0000	0.0040
GARCH(1,1)	2 year	95%	1	248	1	1	0	0.0040	0.0000	0.0040
GJR-GARCH	2 year	95%	1	248	1	1	0	0.0040	0.0000	0.0040
TGARCH	2 year	95%	2	246	2	2	0	0.0081	0.0000	0.0080
ARCH(1,1)	2 year	99%	0	250	0	0	0	0.0000	0.0000	0.0000
GARCH(1,1)	2 year	99%	0	250	0	0	0	0.0000	0.0000	0.0000
GJR-GARCH	2 year	99%	0	250	0	0	0	0.0000	0.0000	0.0000
TGARCH	2 year	99%	0	250	0	0	0	0.0000	0.0000	0.0000

Table A.12: Auxiliary data for the independence test

Appendix B

Programme Code

GARCH model

```
% GARCHrun.m
% input set of log returns
smp = load('hsba_r.txt');
% initial values of model parameters
startParams = [6.87E-06 0.1209  0.8549];

% optimal values of parameters maximizing the likelihood function
parameters_GARCH = GARCHcalibration (smp, startParams);
LLF_GARCH=-garchMaxlikelihood(smp,parameters_GARCH);
```

```
%  GARCHcalibration.m

function parameters = GARCHcalibration (rets, startParams)
% set of parameters of function fminsearch
options=optimset('MaxFunEvals',500,'Maxiter',500,'Display','iter');
    function f = mns_aux(params)
         f = garchMaxlikelihood (rets,params);
    end

% optimal values of parameters maximizing the likelihood function
parameters = fminsearch(@mns_aux, startParams, options);
end
```

```
%garchMaxlikelihood.m (GARCH Likelihood function)
function y = garchMaxlikelihood (rets, startParams)

% initial values of parameters
omega=startParams(1);
alpha=startParams(2);
beta=startParams(3);

% length of the input array
n=length(rets);
if ((omega<0) || (alpha<0) || (beta<0))
    y=intmax;
    return;
end

vart(1)=var(rets); % variance at the day 1, \sigma(1)^2
y=-log(vart(n))-(rets(n)^2/vart(n));

% the values of variance are computing iteratively
for cnt=2:n
   vart(cnt)=omega+alpha*rets(cnt-1)^2+beta*vart(cnt-1);
   y=y-log(vart(cnt))-(rets(n)^2/vart(cnt));
end

y=-y; % the final value of likelihood function
end
```

```
%GARCH_sim_head.m
% it simulates  a given GARCH process

% input values of log returns
smp = load('hsba_log_return_1.txt');

% input values of parameters of GARCH model
Params = [2.83903e-06 0.13817 0.84335];

% length of the input array (period of time)
T = length(smp);
%T = 755;
% simulation
v = GARCHsimulation ( smp, T, Params);

% the plot
plot(v);
```

```
% GARCHsimulation.m
% given parameters of GARCH model and values of log returns
% it returns the array of values of variance

function v = GARCHsimulation ( smp, T, params)

% parameters of GARCH model
omega_ = params (1); alpha_ = params (2); beta_ = params (3);
w(1)=sqrt(var(smp));

for i=2:T
    % the main equation of GARCH model
    w(i) = sqrt(omega_ + alpha_ *smp(i-1)^2+ beta_ *w(i-1)^2);
end
v=w;
end
```

GARCH model augmented with news intensity

```
%head_news.m
% it runs calibration of GARCH model augmented with news intensity

% parameters of a net of initial points of the search by fminsearch
I1=9;
I2=5;
I3=5;

% load input arrays of log returns of the company stocks
smp=load('aviva_r.txt');

% load input arrays of log returns of of FTSE100 index
ftse =load('r_ftse.txt');

% load input arrays of numbers of news per day
news_intensity =load('aviva_news.txt');

% it finds the optimal values of parameters
% maximizing likelihood function
param=GARCHrunALLnews(I1,I2,I3,smp,ftse,news_intensity);
```

```
% GARCHrunALLnews.m
% it chooses the "best" initial points for fminsearch
% by iterating through the net of points
% defined by I1, I2, I3

function z = GARCHrunALLnews (I1,I2,I3, smp_1, ftse, volume)
value_min=0;
for i=9:I1
    for j=0:(I2-1)
        for k=0:(I2-j-1)
            for l=4:I3
            a1=10^(-i);
            a2=j/I2+0.0001;
            a3=k/I2+0.0001;
            a4=10^(-l); a5=0.0001; a6=1.0;
            startParams = [a1 a2 a3 a4 a5 a6];
                parameters = GARCHcalibration_news ...
                    (smp_1, ftse, volume, startParams);
                value = garchMaxlikelihood_news ...
                    (smp_1, ftse, volume, parameters);
                if (value<value_min)
                    value_min=value;
                    z = parameters;
                end;
            end;
        end;
    end;
end;
```

```
% GARCHcalibration_news.m

function parameters = GARCHcalibration_news ...
    (rets, ftse, volume, startParams)

% set of parameters of function fminsearch
% options=...
%optimset('MaxFunEvals',5000,'Maxiter',500, 'Display','iter');

    function f = mns_aux(params)
         f =  garchMaxlikelihood_news(rets, ftse, volume, params);
    end

% optimal values of parameters maximizing the likelihood function
parameters = fminsearch(@mns_aux, startParams);
%parameters = fminsearch(@mns_aux, startParams,options);
end
```

```
% garchMaxlikelihood_news.m
% Likelihood function of GARCH model augmented with news intensity

function y = garchMaxlikelihood_news ...
    (rets, ftse, volume, startParams)

% initial values of parameters
omega=startParams(1);
alpha=startParams(2);
beta=startParams(3);
gamma=startParams(4);
theta1 = startParams(5);
theta2 = startParams(6);

% length of the input array
n=length(rets);
if ((omega<0) || (alpha<0) || (beta<0) || (gamma<0) )
    y=intmax;
    return;
end

% variance sigma(1)
vart(1)=var(rets);
y=-log(vart(1))-(rets(1)^2/vart(1));

% the values of variance are computing iteratively
for cnt=2:n
   vart(cnt)=omega+alpha*(rets(cnt-1)-...
       theta1-theta2*ftse(cnt-1))^2+ ...
       beta*vart(cnt-1)+gamma*volume(cnt);
   y=y-log(vart(cnt))-...
       ((rets(cnt)-theta1-theta2*ftse(cnt))^2/vart(cnt));
end

% the final value of likelihood function
y=-0.5*(y-log(2*pi));
end
```

GARCH model with Jumps (the case of constant jump intensity)

```
%GARCHrun_Jumps_Const.m
% it runs calibration of GARCH model with jumps
% (the case of constant jump intensity)

% load input arrays of log returns of the company stocks
smp=load('hsba_r.txt');

% initial values of model parameters
startParams = [1.17E-06 0.122 0.86 5.119E-03 1.67E-03 0.39];

% it finds the optimal values of parameters
% maximizing likelihood function
parameters = GARCHcalibration_Jumps_Const(smp, startParams);

the final value of likelihood function
LLF_Jumps = -garchMaxlikelihood_Jumps_Const (smp, parameters);

%params_Null=[2.72E-06  0.1232  0.8568  0 0 0];
%LLF_Jumps_Null =- garchMaxlikelihood_Jumps_Const (smp, params_Null);
```

```
% GARCHcalibration_Jumps_Const.m

function parameters = ...
    GARCHcalibration_Jumps_Const (rets, startParams)

% set of parameters of function fminsearch
options=optimset('MaxFunEvals',1000,'Maxiter',1000);

    function f = mns_aux(params)
       f =  garchMaxlikelihood_Jumps_Const(rets, params);
    end

% optimal values of parameters maximizing the likelihood function
parameters = fminsearch(@mns_aux, startParams, options);
end
```

```
% garchMaxlikelihood_Jumps_Const.m
% Likelihood function of GARCH model with Jumps
% (the case of constant jump intensity)

function y = garchMaxlikelihood_Jumps_Const(rets, startParams)
% initial values of parameters
omega=startParams(1);
alpha=startParams(2);
beta=startParams(3);
delta=startParams(4);
theta=startParams(5);
lambda=startParams(6);

n=length(rets); % length of the input array
if ((omega<0) || (alpha<0) || (beta<0) )
    y=intmax;
    return;
end

vart(1)=var(rets); % variance \sigma(1)^2
   f=0;
   fact=1;
   for j=0:20
        f=f+(lambda^j)*exp(-((rets(1)+theta*lambda-theta*j)^2)/...
            (2*(vart(1)+j*delta^2))-lambda)*...
            ((2*pi*(vart(1)+j*delta^2))^(-0.5))*(fact^(-1));
       fact=fact*(j+1);
   end
y=log(f);

% the values of variance are computed iteratively
for cnt=2:n
   vart(cnt)=omega+alpha*(rets(cnt-1))^2+beta*vart(cnt-1);
   f=0;
   fact=1;
   for j=0:20
       f=f+(lambda^j)*exp(-((rets(cnt)+theta*lambda-theta*j)^2)/...
           (2*(vart(cnt)+j*delta^2))-lambda)*...
           ((2*pi*(vart(cnt)+j*delta^2))^(-0.5))*(fact^(-1));
       fact=fact*(j+1);
   end
   y=y+log(f);
end

y=-y; % the final -value of likelihood function
end
```

```
%GARCH_sim_JUMP_head.m
% it simulates a given GARCH-Jumps process

% input values of log returns
smp_1=load('BP_r.txt');

% input values of parameters of GARCH-Jumps model
Params = [7.09e-06 0.15 0.84    1.6202E-02  -1.08E-02   0.12];

% length of the input array (period of time)
T=length(smp_1);

% simulation
v = GARCH_sim_JUMP( smp_1 ,T, Params);

% the plot of simulated variance
plot(v);
```

```
% GARCH_sim_JUMP.m
% given parameters of GARCH-Jumps model and values of log returns
% it returns the array of values of variance

function v = GARCH_sim_JUMP ( smp_1 ,T,params)

% parameters of GARCH-Jumps model
omega_ = params (1); alpha_ = params (2);
beta_ = params (3); delta_=params(4);
theta_=params(5); lambda_=params(6);
w(1)=sqrt(var(smp_1));
for i =2:T
    jumpnb = poissrnd ( lambda_,1 );
    jump =...
        normrnd( theta_ *( jumpnb - lambda_), sqrt(jumpnb)*delta_);
     % the main equation of GARCH model
w(i)  = sqrt(omega_ + alpha_ *(smp_1(i-1)-jump)^2+ beta_ *w(i-1)^2);
end
v=w;
end
```

GARCH model with Jumps

```
%GARCHrun_Jumps.m
% it runs calibration of GARCH model with jumps

% load input arrays of log returns of the company stocks
smp=load('BP_r.txt');

% initial values of model parameters
startParams = [0.000001 0.1 0.9 0.05 0 0 1 0];

% it finds the optimal values of parameters
% maximizing likelihood function
parameters = GARCHcalibration_Jumps(smp, startParams);

% the final value of likelihood function
value_final = garchMaxlikelihood_Jumps (smp, parameters);
```

```
% GARCHcalibration_Jumps.m

function parameters = GARCHcalibration_Jumps (rets, startParams)

% set of parameters of function fminsearch
options=optimset('MaxFunEvals',1000,'Maxiter',1000);

    function f = mns_aux(params)
       f =  garchMaxlikelihood_Jumps(rets, params);
    end

% optimal values of parameters maximizing the likelihood function
parameters = fminsearch(@mns_aux, startParams, options);
end
```

```
% garchMaxlikelihood_Jumps.m
% Likelihood function of GARCH model with Jumps

function y = garchMaxlikelihood_Jumps(rets, startParams)

% initial values of parameters
omega=startParams(1);
alpha=startParams(2);
beta=startParams(3);
delta=startParams(4);
theta=startParams(5);
a=startParams(6);
```

```
b=startParams(7);
c=startParams(8);

% length of the input array
n=length(rets);
if ((omega<0) || (alpha<0) || (beta<0) )
    y=intmax;
    return;
end

lambda=3; %initial value of jump intensity
vart(1)=var(rets); % variance \sigma(1)^2
   f=0;
   fact=1;
   for j=0:10
        f=f+(lambda^j)*exp(-((rets(1)+...
            theta*lambda-theta*j)^2)/(2*(vart(1)+j*delta^2))...
            -lambda)*((2*pi*(vart(1)+j*delta^2))^(-0.5))*...
            (fact^(-1));
       fact=fact*(j+1);
   end
y=log(f);

for cnt=2:n
   vart(cnt)=omega+alpha*(rets(cnt-1))^2+beta*vart(cnt-1);
   z=(f- (exp(-((rets(cnt-1)+...
       theta*lambda)^2/vart(cnt-1))-lambda))*...
       ((2*pi*vart(cnt-1))^(-0.5)))/f;

   % autoregression for computing of the jump intensity
   lambda=a+b*lambda+c*(z-lambda);

   f = 0;
   fact = 1; %factorial
   for j=0:10
       f = f+(lambda^j)*exp(-((rets(cnt)+theta*lambda-theta*j)^2)/...
           (2*(vart(cnt)+j*delta^2))-lambda)*...
           ((2*pi*(vart(cnt)+j*delta^2))^(-0.5))*(fact^(-1));
        fact=fact*j;
   end
   y=y+log(f);
end

% the final value of likelihood function
y=-y;
end
```

GARCH–Jumps Model augmented with news analytics data

```
% GARCHrun_Jumps_News.m
% calibration of GARCH-Jumps model augmented with news intensity

% input set of log returns
smp=load('hsba_r.txt');

% the input file with the number of positive news per day
volume1=load('hsba_ncss_pos.txt');

% the input file with the number of negative news per day
volume2=load('hsba_ncss_neg.txt');

% initial values of model parameters
startParams = [6.87E-06 0.1209  0.8549  0 0 0 0 0];

% optimal values of parameters maximizing the likelihood function
parameters = ...
GARCHcalibration_Jumps_News (smp,volume1,volume2,startParams);

% the final value of likelihood function at the optimal point
value_Jumps_Const = - ...
garchMaxlikelihood_Jumps_News (smp,volume1,volume2,parameters);

%params_Null=[2.21E-06  0.1335...
%0.8278 1.5563E-02  -6.5240E-04 0.0570  0.00158 0.00561];
%LLF_Jumps_Null =- ...
%garchMaxlikelihood_Jumps_News (smp,volume1,volume2,params_Null);
```

```
% GARCHcalibration_Jumps_News.m

function parameters = ...
  GARCHcalibration_Jumps_News (rets, volume1, volume2, startParams)

% set of parameters of function fminsearch
options=optimset('MaxFunEvals',500000,'Maxiter',10000);

    function f = mns_aux(params)
    f =  ...
      garchMaxlikelihood_Jumps_News (rets,volume1,volume2,params);
    end

% optimal values of parameters maximizing the likelihood function
parameters = fminsearch(@mns_aux, startParams, options);
end
```

```
% garchMaxlikelihood_Jumps_News.m
% Likelihood function of ...
% GARCH-Jumps model augmented with news intensity ...
% (the case of constant jump intensity)

function y = ...
  garchMaxlikelihood_Jumps_News (rets,volume1,volume2,startParams)

% initial values of parameters
omega=startParams(1); alpha=startParams(2);
beta=startParams(3); delta=startParams(4);
theta=startParams(5); lambda=startParams(6);
rho1=startParams(7); rho2=startParams(8);

n = length(rets); % length of the input array
if ((omega<0) || (alpha<0) || (beta<0) )
    y=intmax;
    return;
end

vart(1)=var(rets);  % variance \sigma(1)^2
   f=0;
   fact=1;
   for j=0:20
        f=f+(lambda^j)*exp(-((rets(1)+ theta*lambda-theta*j)^2)/...
            (2*(vart(1)+j*delta^2))-lambda)*...
            ((2*pi*(vart(1)+j*delta^2))^(-0.5))*(fact^(-1));
       fact=fact*(j+1);
   end
y=log(f);

for cnt=2:n
    % the values of variance are computed iteratively
   vart(cnt)=omega+alpha*(rets(cnt-1))^2+beta*vart(cnt-1);
   f=0;
   fact=1;
   for j=0:20
       f=f+(lambda+rho1*volume1(cnt)+rho2*volume2(cnt))^j...
          *exp(-((rets(cnt)+theta*(lambda+rho1*volume1(cnt)+rho2*...
           volume2(cnt))-theta*j)^2)/(2*(vart(cnt)+j*delta^2))...
           -(lambda+rho1*volume1(cnt)+rho2*volume2(cnt)))*...
           ((2*pi*(vart(cnt)+j*delta^2))^(-0.5))*(fact^(-1));
       fact=fact*(j+1);
   end
   y=y+log(f);
end
y=-y; % the final value of likelihood function
end
```

Monte Carlo simulation (GARCH model with volume)

```
% input values of log returns
smp=load('lloyd_r.txt');

% load input arrays of volumes of news per day
volume =load('Lloyd_volume.txt');

% input values of parameters of GARCH model
Params = [1.13E-07  0.0928  0.8573  2.26E-13];

% length of the input array (period of time)
T = length(smp);

% the number of MC runs
N=500;

for j=1:N
% simulation
w=MC_GARCH_Volume_simulation ( smp, volume, T, Params);

for s=1:T
e(s) = w(s)*random('normal',0,1);
end

% smp_new=e;
smp_new=e;

startParams = [1.13E-07 0.0928  0.8573  2.26E-13];
startParams_garch = [4.91E-06   0.1230  0.8609];

% optimal values of parameters maximizing the likelihood function
parameters_GARCH_Volume =...
    GARCHcalibration_volume (smp_new, volume, startParams);
omega(j)=parameters_GARCH_Volume(1);
alpha(j)=parameters_GARCH_Volume(2);
beta(j)=parameters_GARCH_Volume(3);
gamma(j)=parameters_GARCH_Volume(4);
LLF_GARCH_Volume(j)=...
-garchMaxlikelihood_volume(smp_new,volume,parameters_GARCH_Volume);
parameters_GARCH = GARCHcalibration (smp_new, startParams_garch);
LLF_GARCH(j)=-garchMaxlikelihood (smp_new, parameters_GARCH);
LL_Ratio(j)=2(LLF_GARCH_Volume(j)-LLF_GARCH(j));
end

mean_omega=mean(omega);
mean_alpha=mean(alpha);
mean_beta=mean(beta);
```

```
mean_gamma=mean(gamma);
mean_LLF=mean(LLF_GARCH_Volume);
var_omega=var(omega);
var_alpha=var(alpha);
var_beta=var(beta);
var_gamma=var(gamma);
var_LLF=var(LLF_GARCH_Volume);
var_LL_Ratio_lloyd=var(LL_Ratio);
mean_LL_Ratio_lloyd=mean(LL_Ratio);

% the plot
hist(LL_Ratio,40);
hist(alpha,20);
hist(beta,20);
hist(gamma,30);
boxplot(alpha);
boxplot(beta);
boxplot(gamma);
pctl = 100*(0:0.05:1);
ypctl = prctile(gamma,pctl);
zpctl = [pctl;ypctl];
zpctl;
```

```
% MC_GARCH_Volume_simulation.m
% given parameters of GARCH model and values of log returns
% it simulates the array of values of variance
function v = MC_GARCH_Volume_simulation ( smp, volume, T, params)

% parameters of GARCH model
omega_ = params (1); alpha_ = params (2);
beta_ = params (3); gamma_=params (4);

w(1)=sqrt(var(smp));

for i=2:T
    % the main equation of GARCH model
  w(i) = sqrt(omega_ + ...
      alpha_ *smp(i-1)^2*(random('normal',0,1))^2+...
        beta_ *w(i-1)^2 + gamma_ *volume(i));
end
v=w;
end
```

Monte Carlo simulation (GARCH model with news intensity)

```
%MC_News_sim_head.m
% Monte Carlo simulation

% load input arrays of log returns of the company stocks
smp=load('aviva_r.txt');

% load input arrays of log returns of of FTSE100 index
ftse =load('r_ftse.txt');

% load input arrays of numbers of news per day
news_intensity =load('aviva_news.txt');

% input values of parameters of GARCH-news model
Params = [4.1157E-11 0.0001 0.9562 2.24E-06 -4.36E-04 1.2086];

% length of the input array (period of time)
T = length(smp);

% the number of MC runs
N=500;

for j=1:N
% simulation
w=MC_GARCH_News_simulation( smp, ftse, news_intensity, T, Params);

for s=1:T
e(s) = w(s)*random('normal',0,1);
end

smp_new=e;

startParams = Params;

% optimal values of parameters maximizing the likelihood function
parameters_GARCH_news = ...
    GARCHcalibration_news (smp_new, ftse, news_intensity, startParams);
parameters_GARCH = [6.87E-06 0.1209 0.8549];

omega(j)=parameters_GARCH(1);
alpha(j)=parameters_GARCH(2);
beta(j)=parameters_GARCH(3);
gamma(j)=parameters_GARCH(1);
theta1(j)=parameters_GARCH(2);
theta2(j)=parameters_GARCH(3);

LLF_GARCH_Volume(j)=-garchMaxlikelihood_news...
```

```
    (smp_new, ftse, news_intensity, parameters_GARCH_news);

LLF_GARCH(j)=-garchMaxlikelihood (smp_new, parameters_GARCH);
LL_Ratio(j)=2(LLF_GARCH_Volume(j)-LLF_GARCH(j));
end

mean_omega_aviva=mean(omega);
mean_alpha_aviva=mean(alpha);
mean_beta_aviva=mean(beta);
mean_gamma_aviva=mean(gamma);
mean_theta1_aviva=mean(theta1);
mean_theta2_aviva=mean(theta2);
mean_LLF_aviva=mean(LLF_GARCH_Volume);
var_omega_aviva=var(omega);
var_alpha_aviva=var(alpha);
var_beta_aviva=var(beta);
var_gamma_aviva=var(gamma);
var_theta1_aviva=var(theta1);
var_theta2_aviva=var(theta2);
var_LLF_aviva=var(LLF_GARCH_Volume);
var_LL_Ratio_aviva=var(LL_Ratio);
mean_LL_Ratio_aviva=mean(LL_Ratio);
```

```
% MC_GARCH_News_simulation.m
% given parameters of GARCH model and values of log returns
% it simulates the array of values of variance

function v = ...
    MC_GARCH_News_simulation ( smp, ftse, volume, T, params)

% parameters of GARCH model
omega_ = params (1); alpha_ = params (2);
beta_ = params (3); gamma_=params (4);
theta1_ = params (5); theta2_ =params (6);

vart(1)=sqrt(var(smp));

for i=2:T
    % the main equation of GARCH model
    vart(i)=sqrt(omega_ +alpha_ *(smp(i-1)-...
     theta1_ - theta2_ *ftse(i-1))^2*(random('normal',0,1))^2+ ...
     beta_ *vart(i-1)^2+gamma_ *volume(i));
end
v=vart;
end
```

Bibliography

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

V. Arago and L. Nieto. Heteroskedasticity in the returns of the mainword stock exchange indices: volume versus garch effects. *International Financial Markets Institute and Money*, 15:271–284, 2005.

R. Bauer and F. Nieuwland. A multiplicative model for volume and volatility. *Applied Mathematical Finance*, 2:135–154, 1995.

T. D. Berry and K. M. Howe. Public information arrival. *Journal of Finance*, 49:1331–1346, 1993.

F. Black. Studies in stock price volatility changes. In *Proceedings of the Business and Economic Statistics Section*, pages 177–181. American Statistical Association, 1976.

T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327, 1986.

T. Bollerslev. Arch modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics*, 52:5–59, 1992.

T. A Bollerslev. Conditionally heteroskedastic time series model for security prices and rates of return data. *Review of Economics and Statistics*, 69:542–547, 1987.

W. H. Chan and J. M. Maheu. Conditional jump dynamics in stock market returns. *Journal of Business and Economic Statistics*, 20(3):377–389, 2002.

Y.-T. Chen and C.-M. Kuan. The pseudo-true score encompassing test for non-nested hypotheses. *Journal of Econometrics*, 106:271–295, 2002.

P. Christoffersen. Evaluating interval forecasts. *International Economic Review*, 39:841–862, 1998.

P. K. Clark. A subordinated stochastich process model with finite variance for speculative prices. *Econometrica*, 41:135–155, 1973.

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

D. R. Cox and D. V Hinkley. *Theoretical Statistics*. Chapman and Hall, 1974.

L. H. Ederington and J. H. Lee. How markets process information: News releases and volatility. *Journal of Finance*, 48:1161–1191, 1993.

R. F. Engle. Autoregressive conditional heteroscedasticity with estimates of variance of united kingdom inflation. *Econometrica*, 50:987–1008, 1982.

R. F. Engle and V. K. Ng. Measuring and testing the impact of news on volatility. *Journal of Finance*, 48:174961777, 1993.

T. Epps and M. Epps. The stochastic dependence of stochastic price changes and transaction volume: Implications for the mixture of distribution hypothesis. *Econometrica*, 44:305–321, 1976.

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

L. R. Glosten, R. Jagannathan, and D. E. Runkle. On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, 48:1779–1801, 1993.

A. R. Hall. *Generalized Method of Moments*. Oxford University Press, 2004.

P. Hansen and A. Lunde. A forecast comparison of volatility models: Does anything beat a garch (1,1)? working paper, Department of Economics, Brown University, 2001.

L. Harris. Transaction data tests of the mixture of distributions hypothesis. *Journal of Financial and Quantitative Analysis*, 22:127–141, 1987.

G. Janssen. Public information arrival and volatility persistence in financial markets. *The European Journal of Finance*, 10:177–197, 2004.

Philippe Jorion. *Value at Risk: The New Benchmark for Managing Financial Risk*. McGraw-Hill, 2 edition, 2001.

P. S. Kalev, W.-M. Liu, P. K. Pham, and E. Jarnecic. Public information arrival and volatility of intraday stock returns. *Journal of banking and Finance*, 28 (6):1447–1467, 2004.

M. Kantardzic. *Data Mining: Concepts, Models, Methods, and Algorithms*. John Wiley & Sons, 2003.

J. M. Karpoff. The relation between price changes and trading volume: A survey. *Journal of Financial and Quantitative Analysis*, 22:109–126, 1987.

S. Kim, N. Shephard, and S. Chib. Stochastic volatility: Likelihood inference and comparison with arch models. *Review of Economic Studies*, 65:361–393, 1998.

P. Kupiec. Techniques for verifying the accuracy of risk management models. *Journal of Derivatives*, 3:73–84, 1995.

C. G. Lamoureax and W. D. Lastrapes. Heteroskedasticity in stock return data: volume versus garch effects. *Journal of Business & Economic Statistics*, 2:253–260, 1990.

C. G. Lamoureax and W. D. Lastrapes. Endogenous trading volume and momentum in stock-return volatility. *Journal of Finance*, 45:221–229, 1991.

L. Le Cam. Maximum likelihood - an introduction. *ISI Review*, 58(2):153–171, 1990.

J. H. Lee and B. W. Brorsen. A Cox-type non-nested test for time series models. *Applied Economics Letters*, 4:765–768, 1997.

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

H. Malmsten. Evaluating exponential garch models. *SSE/EFI Working Paper Series in Economics and Finance, Stockholm School of Economics*, 564, 2004.

M. L. Mitchell and J. H. Mulherin. How markets process information: News releases and volatility. *Journal of Finance*, 49:923–950, 1994.

G. Mitra and L. Mitra, editors. *The Handbook of News Analytics in Finance*. John Wiley & Sons, 2011.

T. Miyakoshi. Arch versus information-based variances: evidence from the tokyo stock market. *Japan and the World Economy*, 14:215–231, 2002.

J. I. Myung and D. J. Navarro. *Encyclopedia of Behavioral Statistics*, volume 2, chapter Information matrix, pages 923–924. Chichester, UK: Wiley, 2005.

M. Najand and K. Yung. A garch examination of the relationship between volume and variability in futures markets. *The Journal of Futures Markets*, 11:613–621, 1991.

D. Nelson. Arch models as diffusion approximations. *Journal of Econometrics*, 45:97–127, 1990.

D. B. Nelson. Conditional heteroscedasticity in asset returns: a new approach. *Econometrica*, 59:347–370, 1991.

V. Ragunathan and A. Peker. Price variability, trading volume and market depth: evidence from the australian futures market. *Applied Financial Economics*, 7:447–454, 1997.

J. Ravi, G. Skoulakis, and Z. Wang. Generalized method of moments: applications in finance. *Journal of Business and Economic Statistics*, 20(4):470–481, 2002.

P. Samuelson. Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6:41–49, 1965.

G. W. Schwert. Stock volatility and the crash of '87. *Review of Financial Studies*, 3(1):77–102, 1990.

E. Sentana. Quadratic arch models. *Review of Economic Studies*, 62:639–661, 1995.

J. L. Sharma, M. Mougoue, and R. Kamath. Heteroscedasticity in stock market indicator return data: volume versus garch effects. *Applied Financial Economics*, 6(1):337–342, 1996.

G.E. Tauchen and M. Pitts. The price variability volume relationship on speculative markets. *Econometrica*, 51:485–505, 1983.

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

R. S. Tsay. *Analysis of Financial Time Series*. Wiley, second edition edition, 2005.

H. White. A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica*, 48(4):817–838, 1980.

J. Xiao, R. D. Brooks, and W.-K. Wong. Garch and volume effects in the australian stock markets. *Annals of Financial Economics*, 5:79–105, 2009.

J. M. Zakoian. Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, 18:931–955, 1994.