

UNIVERSITY OF VAASA

FACULTY OF BUSINESS STUDIES

DEPARTMENT OF ACCOUNTING AND FINANCE

Juha Heikkilä

**FORECASTING CRUDE OIL MARKET VOLATILITY:
TEST OF SYMMETRIC AND ASYMMETRIC GARCH-TYPE MODELS**

Master's Thesis in
Accounting and Finance

Finance

VAASA 2006

TABLE OF CONTENTS		page
FIGURES		5
TABLES		5
ABSTRACT		7
1.	INTRODUCTION TO VOLATILITY MODELLING	9
1.1.	Review of Previous Research	10
1.2.	Purpose, Approach and Hypotheses of the Study	14
1.3.	Organization and the Main Results of the Study	15
2.	TIME SERIES MODELLING AND VOLATILITY	17
2.1.	Random Walk	17
2.2.	Stochastic Processes, Ergodity and Stationarity	18
2.3.	Volatility	20
2.4.	Empirical Findings in Asset Price Volatility	22
2.4.1.	Clustering	22
2.4.2.	Mean Reversion	24
2.4.3.	Asymmetry in Volatility	25
2.4.4.	Exogenous Variables	25
2.4.5.	Tail Probabilities	27
2.4.6.	Conclusions on Modelling Needs	27
3.	VOLATILITY FORECASTING WITH THE ARCH -MODELS	29
3.1.	ARCH (q)	30
3.2.	GARCH (p,q)	32
3.3.	TARCH (p,q)	34
3.4.	EGARCH (p,q)	35
3.5.	Model Comparison	37
4.	DATA AND METHODS	39
4.1.	Data	39
4.1.1.	Oil Market in General	40
4.1.2.	The Brent Crude Market	43

4.2.	Measures for Forecasting Performance	44
4.2.1.	Symmetric Loss Functions	44
4.2.2.	Asymmetric Loss Functions	47
5.	EMPIRICAL TESTS AND FINDINGS	49
5.1.	Empirical Data	49
5.2.	Volatility Model Parameter Estimates	55
5.3.	Empirical Results and Forecast Evaluation	59
6.	CONCLUSIONS	64
	REFERENCES	66

FIGURES	page
Figure 1: Volatility clustering (Alexander 2001: 66).	24
Figure 2: The development of Brent Crude Oil Index from 1/1990 to 10/2005	39
Figure 3: The inverted return series, 4 th January 1993 to 31 st January 1997.	51
Figure 4: The inverted return series, 30 th August 2001 to 28 th September 2005.	51
Figure 5: Descriptive statistics for inverted Brent Crude Oil Index returns from 4 th January 1993 through 31 st January 1997.	53
Figure 6: Descriptive statistics for inverted Brent Crude Oil Index returns from 30 th August 2001 to 28 th September 2005.	53

TABLES

Table 1: The augmented Dickey-Fuller test for first return series.	54
Table 2: The augmented Dickey-Fuller test for second return series.	54
Table 3: Engle ARCH-tests for returns. Both estimation periods as own sample.	55
Table 4: The GARCH model parameter estimates on first period.	56
Table 5: The GARCH model parameter estimates on second period.	57
Table 6: The TARARCH model parameter estimates on first period.	57
Table 7: The TARARCH model parameter estimates on second period.	57
Table 8: The EGARCH model parameter estimates on first period.	58
Table 9: The EGARCH model parameter estimates on second period.	58
Table 10: The first estimation and forecasting period, 1-day forecasts.	59
Table 11: The first estimation and forecasting period, 3-day forecasts.	59
Table 12: The first estimation and forecasting period, 5-day forecasts.	60
Table 13: The second estimation and forecasting period, 1-day forecasts.	60
Table 14: The second estimation and forecasting period, 3-day forecasts.	60
Table 15: The second estimation and forecasting period, 5-day forecasts.	61

UNIVERSITY OF VAASA
Faculty of Business Studies

Author:	Juha Heikkilä	
Topic of the Thesis:	Forecasting Crude Oil Market Volatility: Test of Symmetric and Asymmetric GARCH-type Models	
Name of the Supervisor:	Professor Sami Vähämaa	
Department:	Department of Accounting and Finance	
Major Subject:	Accounting and Finance	
Line:	Finance	
Year of Entering the University:	1997	
Year of Completing the Thesis:	2006	Pages: 74

ABSTRACT

The purpose of this thesis is to compare the predictive power of three different volatility forecasting models on Brent Crude Oil Index data under two different market conditions. The models included are GARCH, TARARCH, and EGARCH. The data covers the period from January 1990 to October 2005. From this overall data two periods of data is extracted both individually representing unique era in the market. First data set measures models functionality during mid 1990's tranquil times and second measures model performance at the era of higher uncertainty in the early 2000's.

Four hypotheses were formed in this study based on the findings in earlier studies. The first hypothesis suggests that the more complex model should generate most accurate forecasts. Second hypothesis inspected if the asymmetric volatility model results more accurate forecasts than the symmetric model. The third hypothesis stated that more volatile period results inferior volatility forecasts. The final hypothesis suggested that the volatility forecasting capability is linked to forecasting horizons length and is decreasing over time.

The empirical tests were concluded by estimating models after two different periods and performing then the forecasting experiment. Each estimation sample was around 4 years and forecasts were constructed for 1-, 3-, and 5-day periods. Forecasting performance of different models is evaluated with five widely used error statistics: the root mean square error (RMSE), mean absolute percentage error (MAPE), the adjusted mean absolute percentage error (AMAPE), logarithmic error (LE), and heteroskedasticity adjusted mean square error (HMSE). Three of four hypotheses were discarded, only third hypothesis was confirmed.

KEYWORDS: volatility forecasting, Brent crude, GARCH, TARARCH, EGARCH

1. INTRODUCTION

Living in the era of ever increasing oil prices has made certain benchmark indices widely followed and traded. At the same time the scientific community is confused over the absolute quantities of oil reserves. One thing is certain: oil is a limited and non-renewable natural reserve. As the commodities are priced by supply, demand and inventory, the price of oil as a scarce commodity is most likely to keep an upward trend in the coming future. The question is: can it be predicted on a some level? This study intends to be more specific: is there any statistical information in path of past returns to help us forecast future volatility? This is a study concerned with forecasting uncertainty in the oil prices. (Geman 2005: 333).

This study takes a closer look at pricing of one of the crude oil markets major indices, namely the North Sea Brent crude oil. Petroleum market is divided between refined and non-refined products, in this study the focus is on statistical pricing behaviour of a non-refined end of the oil commodity market. Since different crude oils differ by the site it is drilled and is such an important commodity, there has been taken some benchmark indices to price other crude qualities in commodities market. In this light the key econometrical forecasting qualities of the Brent crude are interesting for scientist, trader, industrialist or risk manager.

In contemporary finance, volatility has a central role. While at the same time the most basic statistical risk measure and probably the most important one. Statistically, volatility is the asset returns standard deviation in financial time series. Most financial decisions are taken with respect to the volatility that a given asset can exhibit. For example a portfolio manager might rationally want to sell an asset to avoid a portfolio becoming too volatile or a risk manager changes a hedging position of some airliner company to meet changes in oil market volatility. Hence, as financial markets become more and more volatile over the few last decades, it has become ever more important for market participants to continuously follow changes in the asset price process. Due to findings based on copious empirical studies, the historical volatility process seems to hold relevant information for future volatility. The prices seem to be, to some degree, deterministic. Therefore, it appears today to be a rational matter to trade on the basis of the asset return volatility and to manage the

losses that could be suffered for the reason that the volatility is varying over time.

Volatility research can be seen from a couple of aspects. First of all it gives information for finance researchers and therefore is part of their field, but the tools on the other hand are done by the econometricians, there hence, the testing can also be categorized in there. The econometrician's research area falls within triangle between economics, finance and statistics. Even though it cannot be restricted to any of these totally, it has shown its value as a research field contributing the tools for the others. Within the last few decades, one of the most important ideas that econometric time series analysis has contributed to related sciences, itself and real life practitioners alike is the concept of time varying variance in financial market.

This study falls within the subcategory of econometrics known as financial econometrics. This can be defined as the application of statistical techniques to problem solving in finance. These techniques can be useful for testing theories in finance, determining asset prices or returns, testing hypotheses between variables, for financial decision-making, examining the effect on financial markets on changes in economic conditions and forecasting future values of financial variables. (Brooks 2001: 1)

The reason, why volatility (the risk of change) is taken into account so heavily, lies in the fundamental idea of finance: to succeed over risk-free rate of return with the lowest possible risk taken. To achieve this, one must take some level of chances to generate better return. Of course as the prices are thought to fluctuate stochastically, taking chances is all about probabilities. Hence all probabilities have statistically always a distribution. This probability for uncertain event can be therefore estimated, if the underlying distribution and its mean and variance are known.

1.1. Review of Previous Research

This subchapter's purpose is to familiarize reader to main research in the area. Review begins with the portfolio theory, market efficiency and goes to main types of GARCH models.

As noted earlier, the concept of risk is one of the central pieces of finance theory. Studying the connection between risk and variance of financial return has yielded Nobel Prizes in economics 1990 to Markowitz and in 1981 to Tobin, for their work concentrating on portfolio theory 1952 and 1958, respectively. Their studies first time associated risk with the variance of financial return. This was developed further in 1964 by Sharpe, who found that if the market participants behave in that way, then the expected returns should follow his Capital Asset Pricing Model (later CAPM). Only the variances that could not be diversified were rewarded. He also received a Nobel Prize in Economics 1990 for his work developing the CAPM.

In 1976 Fisher Black proposed in his conference article "*Studies in Stock Price Volatility Changes*" to model time-varying nature of asset-return volatility. Until that, volatility was believed to be somewhat constant in financial theory's point of view. He gave three additional suggestions for capabilities to also include in the volatility model. First was that the volatility depends on stock price. This was based on observation that increase in stock price reduces volatility. This logically leads to asymmetry in volatility. Then he noted that volatility tends to return to a long term average. This phenomenon is also known as volatility mean reversion. Finally he found that there are random changes in volatility.

Financial market functions as the valuation system for different sorts of more or less relevant information arriving to markets' knowledge. In 1970 Eugene Fama gave his seminal paper on market efficiency. His cornerstone idea was that if in the market the prices fully reflect available information, it is called efficient. He also categorized three different forms of market efficiency. These are the weak-form, the semi-strong, and the strong-form of market efficiency. Five years earlier Fama (1965) had found clustering behaviour in stock market prices, in latter part of this study there is discussion whether these findings are somewhat inconsistent to Fama's theory of market efficiency or his later second article (1991) on the same theory.

Over ten years before Black's research on volatility, Mandelbrot (1963) and Fama (1965) both reported evidence that large changes are often followed by other large changes and small changes are often followed by small changes in financial time series. Mandelbrot (1963) studied commodity market and Fama

(1965) focused on stock market, both finding asset price clustering. This clustering of large movements and small movements (of either sign) in the assets pricing process was one of the first documented features of the volatility process. Consequently this gave investors a hint of how to model the volatility process. The logical implication of such volatility clustering is that volatility shocks today will influence the expectation of volatility many periods in the future. Or in other words, the volatility as a financial phenomenon has, to a certain level, a memory. This naturally entails that financial market does not absorb relevant information instantly and that volatility itself has value as an information source. Therefore it might be useful to apply a model that allows volatility clustering as it is observed in the market.

The feature of aberrant observations tending to merge in clusters leads naturally to the need of exploiting this feature in order to forecast future volatility. Since volatility is a measure of risk, such forecasts can be useful to evaluate investment strategies. In more particularly, it can be useful for decisions on buying, selling, or more generally, on valuing derivatives or portfolios (Frances 1998; 24 - 25). On the market, the need to make volatility forecasting more accurate for the investor raises the need for taking this kind of phenomena into account. For the observations made by Mandelbrot (1963) and Fama (1965) came a firm theoretical explanation from the findings of Robert F. Engle (1982). He then suggested statistical model, autoregressive conditional heteroscedasticity (later ARCH), for forecasting and modelling clustering in financial time series of time varying volatility. His ingenious idea was to capture the conditional heteroskedasticity of financial returns by assuming that today's variance is a weighted average of past squared unexpected returns. Engle founded ARCH properties in variance estimates of United Kingdom inflation. As a consequence of Engle's work, Tim Bollerslev (1986) suggested a generalisation to ARCH model, and the generalized autoregressive conditional heteroskedasticity (later GARCH) -model was born. This essentially generalizes the purely autoregressive ARCH model to an autoregressive moving average model. The weights on past squared residuals are assumed to decline geometrically at a rate to be estimated from the data.

Daniel Nelson (1991) extended ARCH -model family by his exponential GARCH or EGARCH as it is later known. Central idea behind his extension lies in asset price asymmetry in response to different types of information. Stock

market participants seem to respond more to bad news than in they do when they receive positive information. This asymmetry was already found by Black (1976).

Bollerslev, Chou and Kroner (1992) modelled ARCH in an economics modelling context. Their study summarizes different ARCH family models, the theoretical background and the different uses for different models. They present ARCH modelling to stock- and currency market. Others to survey different types of ARCH -type models are Bollerslev (1994), Engle (2002b), and Engle and Ishida (2002).

Zakoian (1994) and Glosten, Jagannathan and Runkle (1993) developed independently their extension to the ARCH model family which is widely known as the TARARCH or the GJR GARCH. In the GJR GARCH the model name comes as abbreviation of its founder's names and in the TARARCH model name, the letter T comes from word threshold. That describes the model pretty well, since it has build-in threshold mechanism for asset price asymmetry. The models have their motivation from Nelson's (1991) EGARCH model, but have an advantage over this by simpler estimation. Main idea is the same, that positive and negative innovations have different impact on volatility forecast. But now there is a threshold value, which simplifies estimation procedure. The TARARCH and GJR GARCH models differ only by the threshold value in models indicator function, thus the models are interpreted to be the same. (Mills 2000: 137.)

Over the years, the evolution of alternative GARCH -type models has yielded several extensions to the original GARCH model. Some of them continue with asymmetrical path like the Engle's and Ng's (1993) asymmetric GARCH (AGARCH) or nonlinear asymmetric NGARCH -model introduced by Higgins and Bera (1992). Later Duan (1995) has advocated NGARCH volatility model into option pricing framework. Other approaches like Teräsvirta (1996) include solutions to cope with excess kurtosis (which normal GARCH models can't cope) normally seen with high frequency data.

Engle and Patton (2001) scrutinize what makes a good volatility model. They characterize a good volatility model by its ability to capture the commonly held

stylized facts about conditional volatility. They test different types of models from the GARCH family to capture these characteristics.

1.2. Purpose, Approach and Hypotheses of the Study

The purpose of this study is to investigate the forecasting performance of certain econometrical models from the class of generalized autoregressive conditional heteroskedasticity. There is a lot of empirical evidence on the performance of this class of models. The data set used in this study is very interesting; it is the Brent Crude Oil Index data. That is, at least at master's thesis level in finance, quite seldom researched area. Another interesting flavour comes from data period itself, which contains a shock caused by September 11th terrorist attack, the gulf wars, uncertainty trading periods surrounding the last gulf war, and oil market affecting hurricane season in 2005. This naturally raises the question about has it effected volatility models forecasting capabilities? In other words, does more uncertainty automatically yield to poor forecasting results? To test this, the forecasting test is organized in two stages. First one is from more tranquil period during mid 1990's. In the second forecasting period is from 2001 to 2005, which should show if there are any changes in model performance. There will also be taken closer look if the asymmetric set in the GARCH -type volatility model ensures more accurate forecasting in comparison to symmetric model setup. Also the effect of having more complex structure on a model is tested. Affect of having different length in forecasting horizon also taken in closer concern and tested. The precisely stated hypotheses that will be tested in this study are following and were formulated based on earlier studies conducted mainly on a stock market (see Chou (1988); Lumsdaine (1995); Engle & Ng (1993); Taylor (1994); Hagerud (1997)):

1. More complex model yields more accurate forecasts than simpler one.
2. Asymmetric volatility model results more accurate forecasts than the symmetric model.
3. More volatile period results in inferior volatility forecasts.
4. Volatility forecasting capability decreases with longer horizon.

The approach of this study will be positivistic nomothetical approach, which is a typical for this research field. In the traditional nomothetic approach theory is confirmed, or questioned, on a considerable number of statistical observations. A common way is to test modelled hypothesis based on theory with statistical methods on the empirical data. Deduction has central role in this approach. Possible new theories are hypotheses, which are typically sought by induction where observations from the real world give impulses for developing new theories (Salmi & Järvenpää 2000: 263 - 267). In this study both the research problem and hypothesis based on finance theory and econometric theory and earlier research done in these fields. The hypothesis is tested on the Brent Crude Index empirical data and is taken from Thomson Financial DataStream.

From a methodological point of view, the precise goal of this work is to study the forecasting capabilities of different econometrical time series models in the ARCH family.

1.3. Organization and the Main Results of the Study

The study is divided into theoretical and empirical sections. There are six main chapters, including this one. This chapter is followed by chapter where the main corner stones of time series modelling are laid down. This includes getting acquainted with stochastic processes and properties of financial time series. Also the main regular irregularities from the perfect financial market equilibrium are gone through. These include clustering, mean reversion, asymmetry in volatility, exogenous variables, and tail probabilities. However this chapter does not discuss the theoretical part of ARCH -type modelling, describing only the known phenomena and modelling surrounding those in explanatory way.

The third chapter takes a closer look into ARCH -type of modelling from a theoretical point of view and also covers the framework of volatility forecasting. The examination naturally begins with the Robert F. Engle's Nobel winning ARCH -model. Then the focus is shifted to Tim Bollerslev's generalization, the GARCH -model, which introduced easier way to handle lag structure. Then it is a time to look in to Zakoian (1994) and Glosten *et al.* (1993) independently found threshold GARCH structure. After that Daniel B. Nelson's approach to

model asymmetry in asset prices is then introduced. His insight into modelling take into account how asymmetrically good and bad information reflects asset pricing. This can be seen in the financial markets, when the same magnitude of negative information provokes larger shift in asset price than positive information of same magnitude. This anomaly is empirically well documented and is closely discussed already in the second chapter.

The fourth chapter introduces the data of the empirical study, autocorrelation testing, estimation procedure, and forecast evaluation for models used in the empirical test. The purpose of this chapter is to pave the way for the fifth chapter, which contains the empirical tests for the GARCH model, the EGARCH model, and the TARARCH model. Naturally the volatility forecasting rises tallest in this chapter. Making any rational decisions on the forecast performance of any of these models deserves closer scrutiny, it is essential to determine model which perform the best in certain conditions. The last chapter concludes and swiftly discusses study as a whole.

2. TIME SERIES MODELLING AND VOLATILITY

The main purpose of this chapter is to lay the groundwork for ARCH-type modelling and its variations by introducing general play of time series modelling and discussing the vast field of empirically found facts in financial time series. It begins by getting acquainted to basic concepts in time series analysis. A discussion includes introduction to the random walk concept, stochastic processes, and then focus is sifted to time series analysis itself. Then attention is turned to volatility. The last part of this chapter looks into essential empirical findings on volatility in the literature.

2.1. Random Walk

The randomness of financial asset prices is one of the corner stones of the finance theory. The term “random walk” saw its first daylight in a scientific journal *Nature* 1905; see Pearson & Rayleigh (1905). In this case the research problem focused on how to find an optimal way to find a drunk who had been left in the middle of a field. The whole research idea might sound a bit absurd in this context, but it was the one to give a later on the name for concept of how asset prices behave. The solution is to start exactly where the drunk had been placed, because at there is an unbiased estimate of the drunk’s future position, since he will presumably stagger along in an unpredictable and random way. (Mills 1999: 5)

The most natural way to state formally random walk model is as

$$(1). \quad P_t = P_{t-1} + u_t,$$

Where P_t is the asset price observed at the beginning of time t and u_t is an error term. It has zero mean and whose values are independent of each other. The price change $\Delta P_t = P_t - P_{t-1}$, is thus simply u_t , hence independent of past price changes. It is also possible, by successive backwards substitution, to write price P_t as an accumulation of all past errors. (Mills 1999: 5).

Later research like Osborne's (1959) model of Brownian motion implies that equation (1) holds for the logarithms P_t and, further that u_t is drawn from a zero mean normal distribution having constant variance.

The concept of Geometric Brownian Motion gets its impetus from the French mathematician, Louis Bachelier who offered the earliest known analytical valuation for security prices in his mathematics thesis "Théorie de la Speculation" given at the University of Sorbonne (1900, English translation Cootner, 1964). He modelled an elaborate mathematical theory of speculative prices, which he then tested on French government bond prices. His findings were that such prices were consistent with the random walk model. What makes his thesis really remarkable is that he also developed many of the mathematical properties of the Brownian motion which had been thought to have first been derived some years later in physics particularly by, a rather well known gentleman, Albert Einstein. (Mills 1999: 6; Mandelbrot 1989: 86-88).

2.2. Stochastic Processes, Ergodity and Stationarity

Time series is a set of two dimensional observations x_t at time t . This coordinate is in standard time series data either discrete or continuous. In this study, the time series data is discrete. These observations are normally organised in chronological order by discrete time coordinate t . When there is a time series on some specific stock price, then S_t is stock price at some certain moment, t . In empirical time series analysis its common practice to analyse the data after the natural logarithmic transform has been applied. (Frances 2000: 9).

Time series data has most of the time autocorrelation and heteroskedasticity. In this world they are normal phenomena. In fact they are taken into account in estimation and forecasting. Especially autocorrelation has significance in forecasting future values in time series from its past values. This property is modelled by ARIMA -models. If time series can completely be forecasted from past values, it is said to be deterministic. And if some sort of probability distribution is needed, time series is called stochastic. In this study, the time series are stochastic.

When one wishes to analyse a financial time series using formal statistical methods, one must always regard that observed series, (x_1, x_2, \dots, x_T) , as a particular realisation of a underlying stochastic process. A realisation is normally denoted $\{x_t\}_1^T$. While, in general, the stochastic process itself will be the family of random variables $\{X_t\}_{-\infty}^{\infty}$ defined on an appropriate probability space. For this study's purposes it is sufficient to restrict the index set to $T = (-\infty, \infty)$ of the underlying stochastic process to be same as that of the realisation, i.e., $T = (1, T)$, and also to use x_t to denote both the realisation and the underlying stochastic process. (Mills 1999: 8).

If by these conventions the stochastic process is described by T -dimensional probability distribution, so that the relationship between underlying stochastic process and realisation is analogous to that between the population and the sample in classical statistics. The complete specification for the form of the probability distribution will generally be a too ambitious task and it is usual to be content concentrating attention on the first and second moments. If there can also be assumed normality of the probability distribution, this set of expectations would then completely characterise the properties of the stochastic process. The main purpose of these simplifying set of assumptions is that they are made to reduce the number of unknown parameters to more manageable proportions. (Mills 1999: 8-9).

It also has to be emphasised that the procedure of using single realisation to infer the unknown parameters of a joint probability distribution is only valid when the process is ergodic. This roughly means that sample moments for finite stretches of realisation approach their population matching part as the length of the realisation becomes infinite (Mills 1999: 9). In this study the time series is assumed to be ergodic.

Another important simplifying assumption is that of stationarity. This requires process to be a particular state of statistical equilibrium. (Box & Jenkins 1976: 26). If the stochastic process is unaffected by change of its properties time origin it is said to be strictly stationary.

In financial market data, the procedures for return data and for price data are different. When drawn, the basic distinction between stationary and non-stationary time series, it is quite easy to understand. Daily return data on most

financial markets are generated by mean-reverting stationary processes. Actually they are rapidly mean-reverting due to very little autocorrelation in many financial market returns. On the other hand, the statistical methods that apply to return data do not apply to price data. To give an example, correlation and volatility are concepts that only apply to stationary processes. This is because daily (log) price data are assumed to be generated by a non-stationary stochastic process. A good example of such non-stationary processes that are very often applied to prices themselves, or the log prices. (Alexander 2001: 316).

Stationary processes in time series go to higher moments, therefore it is important to take some notice how operators are noted in time series analysis, i.e. addition and multiplication. The first difference operator is defined by

$$(2). \quad \Delta y_t = y_t - y_{t-1}.$$

It is important to note that powers of the first difference operator, such as

$$(3). \quad \Delta^2 y_t = \Delta y_t - \Delta y_{t-1} = y_t - 2y_{t-1} + y_{t-2},$$

should be distinguished from a higher-order difference operators such as

$$(4). \quad \Delta_{12} y_t = y_t - y_{t-12}.$$

Higher-order order differences are used with time series having seasonal components and actually are very useful for this purpose. In (4), for example the 12th difference operator is used to eliminate seasonal effects in monthly data. (Alexander 2001: 316-317).

2.3. Volatility

Volatility in finance is variability of financial asset prices. It is the most common indicator of the level of uncertainty or risk. Volatility is typically expressed in finance as a standard deviation of the random variable. Volatilities are calculated from bond returns, commodity returns, stock returns, interest rates and portfolio market values etc. Expectation of future volatility is in a central role both in practice and finance theory, because of utilize of, and dependence

on volatility forecasts in key financial analysis, investment decision making, and in asset and derivatives pricing. Nowadays risk management, some IFRS - rules and Basel II -framework exploit volatility forecasts increasingly, this even adding value more to accurate volatility forecasting. Mathematically it is a standard deviation of asset return and it is expressed as percents in a year. (Alexander 2001: 4 - 5; Hull 2000: 342).

To forecast, it is naturally of a great importance to know what generates volatility. Numerous factors can be found that cause volatility. In the following subchapter some of these are introduced.

Stock market volatility is generated through the trading process at the market where there is almost as many opinions over the proper value of the financial instrument at trading. Schwert (1989) has studied reasons for volatility changes over time. The analyses included relation of stock volatility with real and nominal macroeconomic volatility, stock trading activity, financial leverage, default risk, and firm profitability using monthly observations 1857 to 1986. He found stock market volatility to be 200% - 300% higher during the Great depression in 1929 - 1939. The macroeconomic series were more volatile during the same period, but could not match the stock market. Also many aggregate economic series such as financial asset returns had greater volatility during recessions. He interpreted it as operating leverage is increasing during recessions.

Schwert (1989: 1145) found weak evidence that volatility of bonds and stock can be forecasted with the help of macroeconomic volatility. When looking into evidence using financial asset prices to predict future macro economic volatility, the results are more promising. Schwert (1989) explains this by concluding that prices of speculative assets absorb quickly new information into prices. Liljeblom and Stenius (1993) tested this question by using Finnish monthly data from the years 1920-1991. They investigated predictive qualities of macroeconomic volatility to predict stock market volatility and vice versa. The conditional return volatility was estimated using two different methods. These were calculated using the GARCH -model and to predict absolute error. The results indicated that changes in stock market volatility did affect to macroeconomic volatility. Liljeblom and Stenius (1995) also repeated their study using the Swedish market data from period 1919 - 1991. The results were

less encouraging. GARCH model did not give any significant results and predicting absolute error didn't either give as promising results as in the Finnish market, though the relationship was there.

Schwert (1989) also showed a relation between trading activity and volatility. A growth in share trading volume and the number of trading days in a month are both positively related to stock volatility.

2.4. Empirical Findings in Asset Price Volatility

In financial and econometric literature, copious reports can be found that describe stylized facts in volatility, central well known observations are gathered into following subchapters.

2.4.1. Clustering

Many financial return series data display volatility clustering. It is one of the first documented features documented in the volatility process of asset returns. This clustering of large moves and small moves of either sign was documented as early as Mandelbrot (1963) and Fama (1965) both of them reported support that large changes in the price of an asset are often followed by other large changes, and small changes are often followed by small changes. This asset price process behaviour has been later supported by numerous other studies, such as Baillie *et al.* (1996), Chou (1988) and Schwert (1989). By these results the implication is that volatility shocks today will influence the expectation of volatility many periods in the future. (Engle & Patton 2001: 242).

This phenomenon where volatility is exhibiting persistence, as it can be put in another way, is a volatility process caused by either the arrival process of news or the market dynamics in response to news. If information comes in clusters, prices or the asset returns may show evidence of ARCH behaviour, even if the market instantaneously and perfectly adjusts to the news. Alternatively the market participants with heterogeneous prior and/or private information may wait or trade some time before the differences of expectations are resolved. (Engle, Ito & Lin 1990: 525 – 526).

Clustering is a market reaction of incoming new information. Intuitively, if on a certain day this information arrives at the market, market participants may react instantaneously by selling or buying assets, whilst after the news has been digested and valued more properly, agents may wish to return to the behaviour before getting the news. When observing time series, from higher frequency to lower frequency, this phenomenon becomes less noticeable. Equity, commodity and foreign exchange markets often exhibit volatility clustering on daily frequency and volatility clustering comes very pronounced in intra-day data. (Alexander 2001: 65; Frances 2000: 24).

In the case of volatility clustering, a rational market participant might want to exploit this in order to forecast future volatility. Being the variable that is used as measurement of risk; such forecasts can be useful to evaluate investment strategies. Furthermore, it can be useful if a model (like GARCH does) takes this account and then use it for decisions on buying or selling options or other derivatives. Time series models that take into account the conditional volatility are often applied to practice and are discussed more in the later part of this study.

A typical example of clustering financial time series is shown in figure 1. Two types of news events are apparent in the figure. Whilst the first event cluster, interpreted by its reaction, this seems to come out of the blue and bear a piece of bad news, the second one is apparently influenced by a scheduled news of a positive nature. The market anticipation, indicated by the growing turbulence, tells it is a scheduled piece of information. Since for a while the conditional mean seems to shift upwards for a while, it is clearly good news for investors. This same logic applies to the first event, but vice versa. The turbulence comes out of nowhere and is shifting the conditional mean clearly to the negative side for a while. (Alexander 2001: 65).

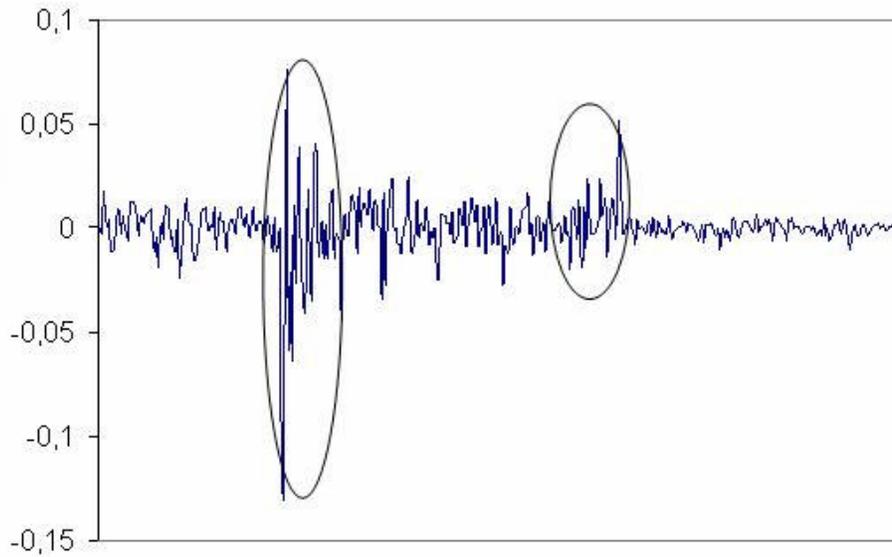


Figure 1. Volatility clustering (Alexander 2001: 66).

2.4.2. Mean Reversion

As volatility clustering implies, volatility comes and goes. After some period of strong volatility patterns, there comes much smoother times in terms of volatility. The stochastic process tends to be near or stabilizes to a long run average value. This is referred as a mean reverting behaviour in the volatility process. It signals that there is a normal long run level for volatility. That is some level where volatility settles after some larger period of turmoil. When scope is a very long volatility prognosis, regardless of the method how prognosis is made, there is a level where all results tend to converge. If there is, and both scholars and practitioners seem to believe there is, a normal level of volatility, mean reversion then implies that current information has no or very little effect on long run forecasts. Also option prices are seen generally consistent with mean reversion. (Alexander 2001: 75; Engle et al. 2001.)

Many papers have documented that the mean reversion pattern i.e. negative autocorrelation is originated by bid-ask effect (see e.g. Miller, Muthuswamy and Whaley 1994; Ederington and Lee 1995; Anderson and Bollerslev 1997).

According to Goodhart and O'hara (1997), the use of higher frequency data appears to underline the evidence of mean reversion having roots in the bid-ask effect. Naturally going lower data frequency should then produce smoother results on mean reversion. Same effect is well documented in studies dealing with commodity data. Schwartz (1997) found in commercial commodity data same mean reverting effect. Crude oil as a commodity belongs to in his classification group of commercial commodities.

2.4.3. Asymmetry in Volatility

The ARCH -type volatility models are built to model volatility shocks. Some are taking into account asymmetry in volatility innovation. This is the case for example EGARCH, but not in symmetrically built models like ARCH and basic GARCH (1,1).

In the world of stock returns it is not realistic to have symmetry in positive and negative shocks. This asymmetry is referred in literature as leverage effect or risk premium effect. The first theory is based on the fact that when stock price falls, the company's debt to equity ratio rises, and thus increasing the volatility of stock returns. The risk premium effect assumes that rising volatility lowers the risk averse investors' interest in that volatile asset. The resulting decline in asset value is followed by the raising volatility as forecast by the news. (Alexander 2001: 68 -69; Engle et al. 2001.)

Fisher Black (1976) found in his article about pricing of commodity options that the returns are negatively correlated with changes in volatility. This naturally means that volatility tends to rise when the market falls and vice versa.

2.4.4. Exogenous Variables

The three phenomena (clustering, mean reversion and asymmetry in volatility) are all univariate characteristics and can be found from time series by looking for information contained in that series' history. No-one believes that financial asset prices evolve without connection to surrounding market. Hence external events (like central bank announcements, OPEC meetings) may and do contain relevant information regarding series volatility. Such evidence has been found

in i.e. Engle, Ito & Lin (1990), and Bollerslev and Melvin (1994), and Nikkinen & Sahlström (2004).

Macroeconomic news, such as Organization of Petrol Exporting Countries (OPEC) meetings, employment, inflation and different price indices do have impact on every asset's volatility. Nikkinen and Sahlström (2004) studied the impact of scheduled macroeconomic announcements and the Federal Open Market Committee's (FOMC) meetings on the implied volatility of S&P 100 index between years 1996 and 2000. The authors investigated the behaviour of implied volatility both around committee's meeting days and announcement dates. Implied volatility was found to have higher levels prior to scheduled announcement and rejoin lower levels after the uncertainty had unveiled. They looked into the days surrounding the employment, the producer price index and the consumer price index. The most notable effect was with employment reports. Furthermore, the FOMC meeting days themselves had significant effect on implied volatility.

For the crude oil markets, the other and competing forms of energy producing give external pressure for volatility. Thus, price of coal or natural gas have their impact on oil market. Different consumption figures have their impact on inventory, supply and demand as well as do the OPEC meetings on production quota. Recently external conditions have influenced greatly in crude market (namely oil futures market) after September 2001 attacks. Then the oil futures plunged after the re-opening of NYMEX, as the market re-calculated after the potential recessionary effects of the World Trade Center attacks. More recently at the beginning of 2003, as second gulf war was ineluctable, a "global insecurity trade" attracted macro investors to go long on commodity options like gold and oil. In this context, exogenous variables like the US White House announcements affected strongly on oil's pricing as investment vehicle. Other sources for exogenous variable are OPEC production quota levels, the inventory levels and obvious changes in demand or supply conditions. Latter conditions change seasonally and are to some extent predictable. (Geman 2005: 201 - 215.)

In commodity volatility modelling literature there is representations that take account of the theory of storage (Kaldor 1939; Working 1949) or the new theory of storage (Williams & Wright 1991). Models containing parameters that take

inventory levels to account should be useful to some extent, when investigating forecasting capability of the volatility model in the context of the crude oil returns. For this Pindyck (2001) develops a theoretical model for how the volatility in principle should affect market variables through the marginal value of storage and through opportunity cost of marginal cost. However, he suggests for petroleum markets that the influence of the changes in volatility for market variables is weak. Market variables do not seem to explain volatility, but as he states it can be forecasted, largely based on its own past values. This study examines only the historical forecasting using information contained within the Brent Crude Oil return series. Thus exogenous variables in forecasting are out of the scope in this study.

2.4.5. Tail Probabilities

The financial theory starts from the assumption that asset returns are normally distributed. Even though Mandelbrot (1963) And Fama (1965) made their seminal contribution to the evidence against normality assumption, it is the easiest way to assume when modelling financial asset returns. Copious studies after them have confirmed their findings.

Engle et al (2001) states that it is a well established fact that the unconditional distribution of asset returns has heavy tails and typically, kurtosis estimates range from 4 to 50. This indicates very extreme non-normality, therefore is a feature that should be incorporated in any volatility model. If the conditional density is normally distributed, then the unconditional density has excess kurtosis due simply to the mixture of Gaussian densities with different volatilities. However there is a little or no reason to assume that the conditional density itself is Gaussian. Actually many volatility models assume that the conditional density is itself fat tailed, thus generating still greater kurtosis in the models unconditional density.

2.4.6. Conclusions on Modelling Needs

When practitioner or scientist takes a look at the volatility modelling, only just presented behaviour in financial market volatility has to be taken care of. Preferably, *a priori*, before actual experiment or using certain model in decision

making in financial markets or at commodity market. At that moment the criteria to be dealt with is consisting all the previous properties of financial market volatility. Volatility model should handle a list of characteristics. These including clustering, mean reversion, asymmetry in volatility, exogenous variables, and changes in tail probabilities. Naturally this leads to growing demand of different qualities to be same time embedded to single model. It has to be same time autoregressive, heteroskedastic, asymmetric, maybe non-linear, possibly multiple equation specification, and possibly usable with non Gaussian distribution specification. The demands for modelling different aspects are obviously great. In the following chapter, the models that are utilized in this experiment to forecast crude oil market volatility are being introduced.

3. VOLATILITY FORECASTING WITH THE ARCH -MODELS

This chapter takes a closer look into ARCH -framework. The introduction of the models follow order from the original ARCH (p) -model, then to generalized representation, the GARCH (p,q) -model. After symmetrically specified models comes the TARCH (p,q) -model, and finally to the most complicated representation, the EGARCH (p,q) -model. The main attention of this chapter is, after the theoretical foundation for time series modelling laid in the last chapter, the conditional variance behaviour within each model. After going through each individual model, follows discussion on these models, their benefits and drawbacks in volatility forecasting. This discussion gives additional information for model selection on different situations.

Generally speaking volatility forecasting is an on going every day activity for risk and portfolio managers as well as many other market participants whether the asset is stock, interest rate or commodity. It is essential to acquire accurate volatility forecasts as swiftly as possible. The econometric challenge in forecasting is to specify how the information is used to forecast the mean and variance of the return, conditional on the past information. For this forecasting effort ARCH and GARCH models are the tool for forecasting asset return variance. Before these models, the primary descriptive tool was the rolling standard deviation. This is obtained by calculating the fixed number of days of the most recent standard deviation observations and letting this “window” to be rolled over time. This assumes that the variance of tomorrow’s return is an equally weighted average of the squared residuals over a pre-specified set of days. The econometricians as well as the practitioners’ point of view this seem unattractive, since all weights are assumed equal. One would think that more recent events would be more relevant holding more information and therefore should bear more weight in the model. Furthermore, the assumption leaves zero weights for observations older than the window specification, which also can leave relevant information out of the return variance forecast. The ARCH and GARCH models let these weights be parameters that are estimated into model. Thus, models following their own specification, allow the data to determine the best weights to use in forecasting. (Engle 2001: 157 - 159.)

3.1. ARCH (q)

The ARCH -class of econometric models was developed by Robert F. Engle in 1982. He received The Bank of Sweden's Prize in Economic Sciences in Memory of Alfred Nobel in 2003 for his work in developing methods for analysing economic time series. This subchapter focuses on his seminal work published in *Econometrica* 1982.

ARCH -models are a class of nonlinear, stationary time series models. In the ARCH process, the conditional variance is estimated to parameters with historical time series values. These processes are stochastic, with the expected value of zero, uncorrelated and whilst the process conditional variance is not constant, the process variance is constant. For these processes, the past observations give information for the coming periods variance forecast.

The stylized facts about observable behaviour of financial time series are well documented. This was presented in a more detailed way in the earlier chapter. In graphical interpretation of the time series, a typical feature is the clustering. As Mandelbrot (1963) found that large (small) change follows a large (small) change of either positive or negative sign, the clustering is reflected in the frequency distribution as fat tails. This results from outliers of both sign and leptokurtosis due to the centring of small changes around the mean. In time series analysis, the family of autoregressive conditional heteroskedasticity models have been developed to account for clustering by explicitly modelling time variation in the second and higher moments of the conditional frequency distribution, which is assumed to be normal. The assumption of the normal density function is convenient in that it enables probability statements about the conditional variance.

In the ARCH models heteroskedasticity is treated as an intrinsic quality of data. This of course has to be modelled, in contrast to econometric analysis before ARCH -type models, the heteroskedasticity was interpreted as a sign of model misspecification. In other words, the main source for conditional variance is not seen coming from past values, but exogenous variables. This leads logically to incorporating the exogenous variable into the model itself. ARCH and GARCH models consider heteroskedasticity as a variance to be modelled. Way the

specification in model does that varies over different ARCH –family model’s specifications. (Ahlstedt 1998: 28; Engle 2001: 157).

The ARCH approach has been used not only in modelling the time series of key financial return series, such as the changes in the foreign exchange rates, interest rates, commodities and stock prices, but also to test financial theories by introducing concept of time-variation to modelling. (Ahlstedt 1998: 28)

When turning the focus to modelling itself the ARCH (q) model can be specified as follows. Following the seminal paper of Robert F. Engle, but using this study’s notation on the model, the conditional variance of a discrete time stochastic process u_t may be denoted σ_t^2 . Which is written as:

$$(5). \quad \sigma_t^2 = \text{var}(u_t | u_{t-1}, u_{t-2}, \dots) = E[(u_t - E(u_t))^2 | u_{t-1}, u_{t-2}, \dots].$$

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

$$(6). \quad \sigma_t^2 = \text{var}(u_t | u_{t-1}, u_{t-2}, \dots) = E[u_t^2 | u_{t-1}, u_{t-2}, \dots].$$

The latter equation states that the conditional variance of a zero mean and normally distributed random variable u_t is equal to the conditional expected value of the square of u_t . The autocorrelation of volatility is modelled in the ARCH (q) model by allowing the conditional variance of the error term, σ_t^2 , to depend on the immediately previous value of the squared error:

$$(7). \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i u_{t-i}^2.$$

where σ_t^2 is a time-varying positive and measurable function of the information set at time $t-i$. By the *IID* assumption, u_t is serially uncorrelated with zero mean. As the $\alpha_0 > 0$ and $\alpha_i > 0$ for all i , non-negative constraining for the parameter values it is necessary to ensure that the conditional variance stays always positive and may change over time. (Ahlstedt 1998: 28–29; Brooks 2002: 445–448).

The variance is always stated as a linear function of past squared values of order q in the ARCH (q) model. From the parameterization of variance in ARCH model, the stochastic process founded in the ARCH framework is not a random

walk but is a martingale. This rules out correlation but allows for dependence in u_t . The time-dependent formula for the conditional variance captures the tendency toward volatility clustering that is often found in financial data. The α_i parameters measure the persistence of shocks in the model. (Engle 1982: 287; Ahlstedt 1998: 28–29; Brooks 2002: 445 – 448)

The order of the ARCH (q) process can be based on model selection tests, such as those which are based on the autocorrelation function of the squared residuals. Many applications of the linear ARCH model have to use a long lag structure. In this case, a normally large order in q leads into a collision course with no negativity constraints on the α_i 's. Fortunately Tim Bollerslev found a solution for this problem in the ARCH -framework by introducing in 1986 his generalized version of ARCH, the GARCH (p,q) model.

3.2. GARCH (p,q)

In 1986 Engle's student Tim Bollerslev introduced a new solution for long lag structures in ARCH -type modelling, with his GARCH -model. It solved a problem, often faced in ARCH modelling, that is when trying to get a good variance forecast the p is grows too large and causes problems in the nonnegative assumption of the model. His model is also capable, of allowing changes in conditional mean, describing phenomena often seen in empirical data called mean reversion. In Bollerslev's (1986) GARCH (p,q) -model the σ_t^2 follows the process giving alternative and more flexible lag structure

$$(8). \quad y_t = u_t, \quad u_t \sim N(0, \sigma_t^2)$$

$$(9). \quad \sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \dots + \alpha_q u_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_p \sigma_p^2$$

$$= \alpha_0 + \sum_{i=1}^q \alpha_i u_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j},$$

where $\alpha_0 > 0$, $\alpha_i > 0$ and $\beta_i > 0$ for all i . The conditional variance depends linearly on the past behaviour of the squared values in an autoregressive AR(q) process and on past values of the conditional variance itself a moving average MA(p) process. The sum of parameters α_i and β_j dictates the persistence of shocks in the model. (Wang 2003: 36; Brooks 2002 452 – 455.)

If the equation (9) p is set to zero, the model naturally changes to an ARCH (q) - model and by repeated substitution it can be shown that the GARCH model is simply an infinite-order ARCH model with exponentially decaying weights for large lags. A high-order ARCH can therefore be substituted by a low-order GARCH model, thus diminishing the problem of estimating many parameters subject to nonnegative constraints. The GARCH (1,1) corresponds to a high-order ARCH (q) of the form

$$(10). \quad \sigma_i^2 = \frac{\alpha_0}{(1 - \beta_1)} + \alpha_1 \sum_{j=1}^{\infty} \beta_1^{j-1} u_{i-j}^2.$$

The conditional variance equation (10) can be interpreted as a one-step-ahead forecast expression. With time series testing procedures, the finding of the optimal parameter values for p and q can be facilitated. The GARCH (1,1) model has proven to be an adequate representation for most financial time series, at least in real world applications. (Ahlstedt 1998: 29 - 30; Brooks 2002: 452 - 455).

In GARCH models, there are also conditions for stationarity to be met. As the name of the model suggests, the variances specified are conditional. As the processes possess a finite variance, the following condition must be met:

$$(11). \quad \sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1.$$

In the most commonly used GARCH (1,1) models, the condition goes simply $\alpha_1 + \beta_1 < 1$. Empirical findings in copious studies suggest that many financial time series have persistent volatility, that is, the sum of α_1 and β_1 is close to being one. This aggregated sum of alpha and beta near unity leads to so-called integrated GARCH or IGARCH as the process no longer holds covariance stationarity. According to Nelson (1990) this still leaves the standard asymptotically based inference procedures generally valid, holding ergodicity or being strictly stationary. (Wang 2003: 36 - 37.)

In other words, an intuitive interpretation of the GARCH (1,1) model is easy to comprehend. There are three components, the GARCH constant term ω (or α_0

as noted in the general form), the GARCH error coefficient α , and the GARCH lag coefficient β . Then the symmetric GARCH (1,1) goes as:

$$(12). \quad \begin{aligned} \sigma_t^2 &= \omega + \alpha u_{t-1}^2 + \beta \sigma_{t-1}^2 \\ \omega > 0, \alpha, \beta &\geq 0 \end{aligned}$$

The GARCH forecast variance is a weighted average of three different variance forecasts. One is a constant variance that corresponds to the long run average. The second is the forecast that was made in the previous period. The third is the new information that was not available when the previous forecast was made. This could be viewed as a variance forecast based on one period of information. The weights on these three forecasts determine how fast the variance changes with new information and how fast it reverts to its long run mean. When the model is seen this way, it reveals the simple ingeniousness behind the GARCH specification. (Alexander 2001: 72 - 75.)

3.3. TARCH (p,q)

The threshold GARCH model or GJR model as it is also known, the latter name coming from the initials of the founders Glosten, Jagannathan and Runkle (1993) the model was also independently founded by Zakoian (1994). The model can be seen as simplified version of EGARCH or a simple GARCH with asymmetric leverage effect variable in its indicator function. Since EGARCH is technically difficult as it involves highly non-linear algorithms to model news impact curve. Though computing power ever increases, when time it self is a factor, simpler estimation has advantages when determining volatility forecasts or doing value at risk analysis etc. The TARCH model enjoys a much simpler estimation method, though not as elegant as, the EGARCH. (Wang 2003: 38-39).

The GJR GARCH and TARCH are in fact the same model. In their article Glosten *et al.* (1993) specify the GJR GARCH indicator functions leverage term $\gamma=2$ and Zakoian (1994) specifies it in TARCH to be $\gamma=1$. The models are otherwise similar. These threshold coefficients allow quadratic response of volatility to news but different coefficients for good and bad news. Nonetheless

it maintains the assertion that the minimum volatility will result when there is no news. (Mills 2000: 137).

The TARCh model is a simple extension of the basic GARCH with an additional term added to take into account for possible asymmetries in financial return series behaviour. The conditional variance is now given by

$$(13). \quad \sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma u_{t-1}^2 I_{t-1}$$

Where $I_{t-1} = 1$ if $u_{t-1} < 0$
 $= 0$ otherwise

For the leverage effect $\gamma = 0$. The condition for non-negativity is now $\alpha_0 \geq 0$, $\alpha_1 \geq 0$, $\beta \geq 0$, and $\alpha_1 + \gamma \geq 0$. So, γ catches asymmetry in response of conditional volatility to shocks in a manner that imposes prior intuition for a positive shock and a negative shock of the same magnitude, future volatility is at least the same or higher, when the sign is negative. This may make sense in many circumstances but not always, like it is a case in commodity markets. (Mills 2000: 137; Brooks 2002: 469 – 470; Wang 2003: 38 – 39.)

3.4. EGARCH (p,q)

Daniel B. Nelson introduced his exponential GARCH model 1990, to capture the asymmetric impact of shocks on the conditional variance. This asymmetry is found particularly in share price data and in inverted form in commodity data. Negative innovations, the negative news, as known in real world, increase volatility more than positive innovations. The linear GARCH model is hence unable to capture this dynamic pattern, since the sign of the shocks plays no role in the symmetric conditional variance model. This asymmetry in mind he embedded asymmetric news impact curve into the EGARCH model. In the EGARCH, the leverage effects are modelled in the conditional variance as an asymmetric function of past innovations. There is numerous ways to express this conditional variance equation, one possible specification is given by:

$$(14). \quad \ln(\sigma_t^2) = \omega + \beta \ln(\sigma_{t-1}^2) + \gamma \frac{u_{t-1}}{\sqrt{\sigma_{t-1}^2}} + \alpha \left[\frac{|u_{t-1}|}{\sqrt{\sigma_{t-1}^2}} - \sqrt{\frac{2}{\pi}} \right].$$

The EGARCH specification has many advantages over the vanilla GARCH model. First, since the $\log(\sigma_t^2)$ is modelled, thus even if negative parameters, σ_t^2 will be positive. Hence, eliminating the need for artificially implying non-negativity constraints on models parameters. Second, the EGARCH model accepts asymmetries, since if the relationship between volatility and returns is negative the news impact curve $\gamma \frac{u_{t-1}}{\sqrt{\sigma_{t-1}^2}}$, will be negative. (Mills 2000: 137; Brooks 2002: 470 - 471.)

It is important to note, that in the original formulation of Nelson (1991) assumed a Generalized Error Distribution (GED) structure for errors. This is a very broad family of distributions that can be used for many types of time series. However, rather than using GED, almost all applications of EGARCH employ conditionally normal errors, the reason being mostly computational ease and intuitive interpretation. An on going study being conducted with the EViews 5 -software package, this point is therefore relevant. (Brooks 2002: 471.)

The asymmetric response parameter or leverage parameter is expected to be positively signed in most empirical cases. The negative sign increases future volatility or uncertainty, while positive shock eases the future uncertainty. This is the feature that is in contrast to the basic GARCH model, where shocks of either sign have the same effect on future uncertainty, which is future volatility. In this model, the conditional variance depends on both the magnitudes and signs of past shocks in the process path. In economic analysis, financial markets and corporate finance, a negative shock usually implies bad news, making the future more uncertain and therefore laying more return expectations in risk conscious investors' minds.

3.5. Model Comparison

When finding a suitable model to volatility forecasting, it is useful to take into account all previous stylized facts in financial time series. Also general knowledge in return series is useful. Many econometrics text books still advice to look at visually at the return series before making up one's mind. Also other knowledge of any possible phenomena mentioned earlier gives pretty good indication for the model selection.

The traditional ARCH model has couple of disadvantages. Naturally, the question arises how the value of q should be decided? One approach is to use a likelihood ratio test. Secondly the number of lags of the squared error (q) that are required to capture all the dependence in the conditional variance could grow very large. This would lead into a large conditional variance model that is not parsimonious. Engle (1982) circumvented this problem by specifying an arbitrary linearly declining lag length on his ARCH (4) model. The last challenge, when operating with ARCH model is that non-negativity constraints can be violated. When there are an increasing number of parameters in a conditional variance equation, then it is more and more likely that one or more of them will have negative estimated values. The GARCH model is the natural extension for ARCH (q) model. They can be estimated in lesser lag structures, GARCH (1,1) is used normally in literature, and therefore they are more parsimonious models. (Brooks 2002: 452.)

The GARCH model has advantages of being relatively easy to estimate and has rather robust coefficients. However the constant parameter (as well as the other two) is especially sensitive to the data used when estimating from historical data. Thus choice of estimation data will strongly affect the current volatility forecasts, particularly long-term volatility estimates will be influenced by the inclusion of the volatile period in the historic data. The problem in the choice of data is always a trade of in statistical forecasting, and is not limiting only into the symmetrical GARCH model, but also to the TARARCH and the EGARCH models. (Alexander 2001: 75; 84 - 85.)

The TARARCH and EGARCH as asymmetrical models take in account the asymmetry in asset price volatility. The TARARCH model is a simpler approach,

and EGARCH being more complex model. In statistical volatility forecasting they work fine as long as the asset return has anticipated asymmetry in its volatility process. For this study this is one very interesting quality to look for. As the price fall is bad news for equity shareholders, it is generally the opposite in the commodity markets where the price falls are the good news and the price rises are the bad news. Due to this characteristic at the commodity market, the return series is inverted for the empirical test. This way, by changing the sign, the asymmetric models should perform as they were designed. (Alexander 2001: 31.)

4. DATA AND METHODS

The following chapters familiarize the reader with the data used in this study, oil market in general, and to the Brent crude market. After getting from general to the more specific picture on data, commodity markets and the crude oil, it is time to go through the test statistics available on model's forecasts for testing their forecast accuracy.

4.1. Data

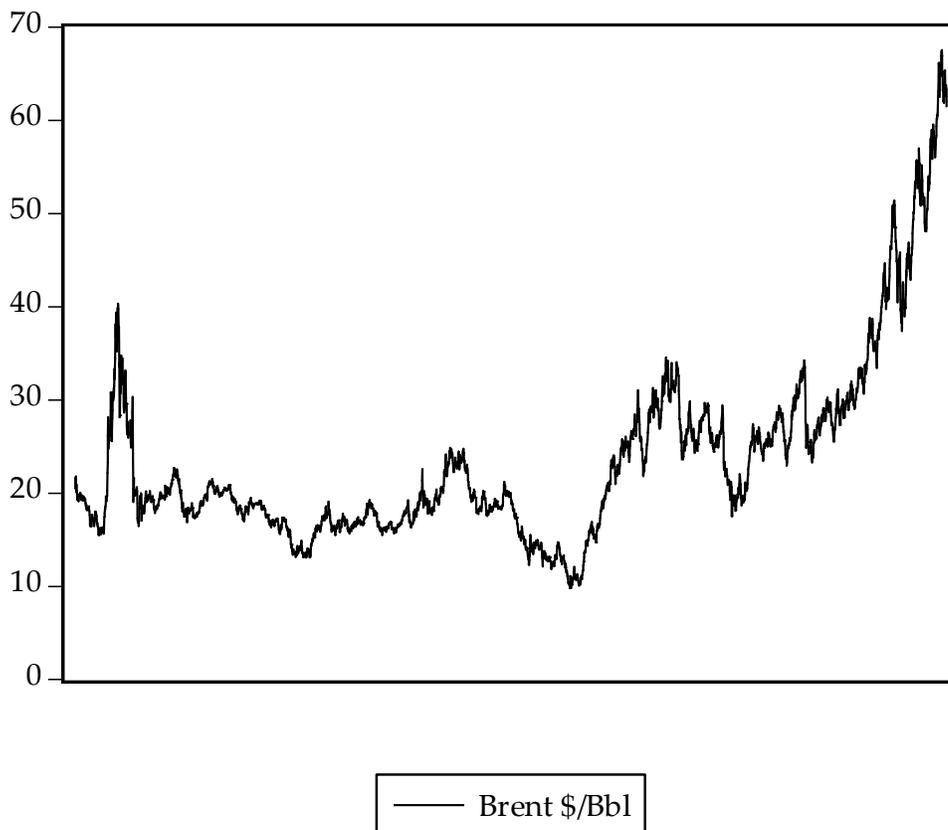


Figure 2. The development of Brent Crude Oil Index from 1/1990 to 10/2005.

The data (seen above in figure 2.) in this study consists 4111 daily price observations between January 2nd 1990 and October 5th 2005. Naturally this

leads to 4110 daily log-return observations from January 3rd 1990 until the end of the data period. The forecasting test is constructed by using two different data periods for each model. First data period is from more tranquil period at mid ninety nineties and the second one test forecasting capabilities after the September 11th 2001 terrorist attacks. This not only should validate if there is superior model, but it should also give forensic evidence about how well the volatility models handle different market conditions. The source for data was Thomson Financial DataStream and it was provided by the University of Vaasa's department of Accounting and Finance.

The log-return data was inverted. This maneuver is needed to set data ready for asymmetric volatility models such as the EGARCH and the TARARCH. They are designed to handle "normal" financial market leverage effect and as the same phenomena appear inverted at the commodity markets, it is logical to invert return series. Otherwise this operation does not have effect on modeling and estimation, only the sign changes in return series. Other important detail is that this time series was not cleaned from outliers. In many cases the outliers are cleaned out of the study if the shocks are not important for the research problem. In this study they are left to preserve two very different market conditions.

4.1.1. Oil Market in General

Oil markets as well as other commodity market categories are viewed as a separate asset class to other "normal" investment goods. This they are, because they cannot be priced in terms of the net present value. A bond, of any kind, is priced as the discounted expectation of future coupon and principal payments. Logic stays the same on pricing of a stock, when one sees the dividends as future cash flows. In other corner stone pricing models – such as the CAPM (see Markowitz 1952) which states the investor is rewarded for the time value of money put upfront to purchase the stock and risk taken. This cannot be extended to commodities, which are priced by supply, demand and inventory. Besides discussed differences in methodological approaches, commodities are seen as a rather distinct asset class due to their counter-cyclic nature. This can be seen from a commodity spot-prices and futures prices for the last 45 years – they have out paced the inflation whole time. In the first years of 21st century, at the era of historically low interest rates and rather poorly profit-making stock

markets, the commodity markets have performed well thus rising interest and turning the eyes of the financial community more and more to new asset classes like commodities. In the past, the commodities were essentially seen as a protection against inflation, now they are recognized as an asset class in its own right, providing not only the diversification benefits to the portfolio of stocks and bonds, but also high returns. (Geman 2005: 333).

Over the past decades the oil market has become the biggest commodity market in the world. The years have attracted more investors to markets than traditional oil traders. Hence it has grown also into a vibrant financial market, with participants from large international financial institutions and funds to physical oil traders and oil refining companies. The market so has risen not only to accommodate basic inventory keeping in the whole refining process and the related fluctuations in supply and demand conditions, but is also sufficient in trading volumes for hedging and speculating. (Geman 2005: 201).

In the 1970's most of the oil was refined by the same producer who drilled it in the first place. That was until the nationalization by oil field hosting countries divided exploration & production and refining operations. Also then, the ambitions for internationalization of the oil market have lead to a situation where the original producer of crude oil seldom refines it. Now the own refining percentage is small compared to the company drilling volumes. Companies trade oil outside their own supply network if they find better opportunities existing in the market. Just in time (JIT) philosophy is widely adopted by the major players. It is important to understand the dynamics of the oil market; the oil is actually physically traded twice. The first stage is, when its refinery feed stock as crude oil and the second time it's traded as a finished product. The study at hand focuses on oil market volatility forecasting in the crude oil market. (Geman 2005: 201 - 202).

At the same time as the financial market -type market conditions for crude oil was developing, the market participants needed some tools for hedging the crude price and some base indices for benchmarks. The latter demand rises from the non-standard nature of crude oil as a commodity. Other crude qualities need to have the benchmark to be priced in relation to that. This development and grown financial market interest has gotten a full set of derivative instruments (futures, forwards, options and swaps) into the market.

Some of the instruments are standardized, other are traded OTC. The most important crude oil derivatives are a futures contract on light sweet crude, WTI (West Texas Intermediate) quoted on the New York Mercantile Exchange (NYMEX) and a futures contract on North Sea Brent Crude quoted electronically at Intercontinental Exchange (ICE). Naturally derivatives attract even more financial market investors trading oil as an asset or hedging their position if their main business is in some oil dependent industry (for example airlines or some process industry areas). These futures contracts on WTI and Brent are the most important on trading all the other crude oil qualities. In real life this is a simplifying routine, otherwise every crude quality then should be priced individually. Indices give clarity and simplicity for the market. Oil price has effects on other markets. The electricity markets in Europe follow crude prices, as well as the international coal market prices. The price of long-term contracts in Europe of delivering Liquefied Natural Gas (LNG) are normally tied to reported prices of fuel oil and gasoil (also known as heating oil) in the Amsterdam-Rotterdam-Antwerp refining hub. The LNG produced in Asia, Africa, Central America and the Middle East is most of the times indexed on baskets of crude oil. In many ways, the price of major crude oil indices are setting pace for the global energy industry, hence it is something worth investigating. (Geman 2005: 201 - 203).

The crude oil markets are highly liquid, global and volatile. At the same time when there are physical market participants trading to keep the wheels turning a 84,6 million barrel per day market, there is an ever increasing flock of investors, speculating on their commodity portfolios (Geman 2005: 204; OPEC 2006). Number of oil investing commodity funds are increasing, thus popularizing the use of crude indices as pricing information benchmarks, like the one on North Sea Brent crude.

The physical market for crude oil depends on the specific grade of crude oil. There are around 400 grades traded world wide. The market value for a crude grade depends essentially on two factors. The first is how many yields of products (butane, propane, gasoline, jet/kerosene, heating oil, and fuel oil) can be extracted on refining process. Quantity and appearance of these yields are directly related to the density of the grade of crude oil under consideration. The second criterion is the amount of energy that must be spent in refinery treating units to remove the sulphur contained in the crude in order to meet tight

quality specifications for refined products imposed by most consuming nations. So when the heavy crude grades are more viscous, contain less of valuable gasoline and are harder to exploit, the light crude oils are very fluid, easy to refine and are rich in gasoline. Naturally the latter light grades are more valuable. The sulphur content in crude grade is referred to as sweet and sour. The sweet qualities contain rather low amounts, less than a percent of its weight of sulphur and sour qualities contain more. Sour qualities are less desirable, since when burning hydrocarbons, the sulphur is turning into sulphur dioxide: a gas that pollutes the air and contributes to acid rain. In general, light oils tend to be sweet, whereas heavy oils tend to be sour. (Geman 2005: 204 – 207).

4.1.2. The Brent Crude Market

The Brent field is one of the older fields of the UK Continental Shelf. Originally, its crude stream was enough to maintain a very active spot market. Its fragmented ownership structure made it a suitable physical basis for a forward paper market. At one point, thirty companies had an equity share in the stream. The field has been drained over time, causing pressure to attach other streams to maintain the quantity of oil in the physical market. In July 1990, the Ninian system output was combined to the Brent. Later in 2004 an index called the Brent BFO was created. It holds additionally streams of Forties and Osberg fields. The reason for combining more streams lies in sufficiency of streams to serve the spot market. There weren't enough spot transactions to keep up the daily price for the Brent on its own. (Geman 2005: 206, 210).

The physical commodity called Brent crude is then in fact a blend of neighbouring oil fields. The Brent blend is a light sweet North Sea crude oil that serves as an international benchmark grade. The significance of the Brent BFO in terms of physical production in international oil trade is small. The Brent blend production runs approximately 500,000 barrels a day. The world daily demand for crude oil was during first quarter 2006 million barrel (OPEC 2006). Most is refined in Northwest Europe, but significant volumes move to the Mexican Gulf and The U.S. East Coast and to the Mediterranean from the shipping terminal at Sullom Voe in the Shetland Islands. It is generally priced FOB (Free On Board). The producing companies trade most of the volume on a spot basis with virtually no formal term contracts. The Brent futures contract markets are based on this spot market. It is now possible to trade these

contracts in the Intercontinental Exchange (ICE), and in NYMEX Europe exchange.

4.2. Measures for Forecasting Performance

For copious different volatility forecasting models, there naturally has to be some measure to examine their performance in relation to the real world volatility. Comparing forecasting performance is one of the most important aspects of any forecasting performance comparisons. Poon and Granger (2003) have made an excellent review article on a volatility forecasting. Naturally their study contains a discussion about measures for a forecasting performance. In order to test this capability, different researchers have approached it in different ways (see e.g. Brailsford & Faff 1996; McMillan, Speigh & Gwilym 2000). The testing procedure, where evaluation methods are used is called a predictive test. This means that the model performance is evaluated by comparing values forecasted by a certain model to the actual data from time series (Alexander 2001: 445). Another role for these estimation measures or error measures as also known is when someone is calibrating or refining a model in order to make ever more accurate forecasts for a set of time series (Armstrong & Collopy 1992).

In this study tests are organized into two categories. These classes are symmetric loss functions and asymmetric loss functions.

4.2.1. Symmetric Loss Functions

If the error measure gives a equal weight to the under and over predictions of the same level of volatility, it is called symmetric loss function. Though used extensively in the practical world of financial market decision making, they have had their share of criticism. The most common symmetric measures are the mean error (ME), the root mean squared error (RMSE), the mean absolute error (MAE), the mean absolute percentage error (MAPE) and Theil -U. (Alexander 2001: 445; Franses 2000: 64 - 65).

The mean error (ME) is pretty self explanatory and is suitable for giving a general guide for interpretation if there is over or under prediction apparent on the forecast in relation to actual series.

$$(15). \quad ME = \frac{1}{T} \sum_{i=1}^T (\hat{\sigma}_i^2 - \sigma_i^2).$$

Where $\hat{\sigma}_i^2$ is the forecasted volatility value of the actual volatility value, that is σ_i^2 . T is the number of periods.

One widespread accuracy measure is the root mean square error or RMSE as it is known. It also is quite self explanatory by its name; it is the square root of the mean of the squared prediction errors. It is defined as:

$$(16). \quad RMSE = \sqrt{\frac{1}{T} \sum_{i=1}^T (\hat{\sigma}_i^2 - \sigma_i^2)^2}$$

The notation follows previous equation (Pindyck & Rubinfeld 1998: 210). In Armstrong et. al. (1992) RMSE is criticised to have low reliability. Alexander (2001: 122 - 123) does not see the use of RMSE between forecast and realised volatility free of problems. She warns that the RMSE test yield normally poor results, because although the expectation of the squared return is the variance, there is a large standard error around this expectation. That is, the squared errors will jump about excessively while the variance forecasts remain more stable. The only justification for using the RMSE between a forecast and the ex-post realized volatility is accordingly to Alexander (2001: 123), which is a simple distance measure. Yu (2000) finds RMSE not to be invariant to scale transformations. The RMSE is also symmetric, so it penalises over and under forecasts the same way.

Similar discussion follow in the literature (for example Brailsford et. al.1996; Brooks 1998) for other popular measures like the mean absolute error (MAE) and the mean absolute percentage error (MAPE). The mean absolute percentage error can be seen as the average prediction error as it is the average of the difference between predicted and actual value. Following previous notation:

$$(17). \quad MAE = \frac{1}{T} \sum_{i=1}^T |\hat{\sigma}_i^2 - \sigma_i^2|$$

The MAPE is probably the most widely used unit free measure. It is seen in original form in equation (4.4):

$$(18). \quad MAPE = \frac{1}{T} \sum_{i=1}^T \left| \frac{(\hat{\sigma}_i^2 - \sigma_i^2)}{\sigma_i^2} \right|.$$

It is the average of the absolute values of errors expressed in percentage terms. Madrakis (1993: 528) liked the MAPE as the best relative measure that incorporates needed characteristics among the various accuracy criteria. In this brief article, he raises four characteristics in MAPE which need to be taken into account in interpreting and using MAPE as a forecast evaluation method. One of these challenges with MAPE can be easily corrected and these notations are taken into account in this study. Madrakis (1993) notes that equal errors above the actual value result in a greater APE (Absolute Percentage Error) those bellow the actual value. This error can easily be corrected by dividing the error between actual and forecast by the average of both as seen in equation (4.5), thus creating a symmetric version of MAPE. Again continuing with the same notation, the Adjusted MAPE (AMAPE):

$$(19). \quad AMAPE = \frac{1}{T} \sum_{i=1}^T \left| (\hat{\sigma}_i^2 - \sigma_i^2) / \left(\frac{\hat{\sigma}_i^2 + \sigma_i^2}{2} \right) \right|$$

Another widely used method is Theil's U -statistic (1966). There the error of forecast is standardised by the error from benchmark forecast. There is obtained typically a simple model, such as a naive or random walk. In this study, the naive forecast is used as benchmark and so it is assumed to be martingale.

$$(20). \quad Theil - U = \frac{\sum_{i=1}^T (\hat{\sigma}_i^2 - \sigma_i^2)^2}{\sum_{i=1}^T (\hat{\sigma}_{BM(i)}^2 - \sigma_i^2)^2}.$$

There the $\sigma_{BM(i)}^2$ is the benchmark forecast and it is used here to remove the effect of any scalar transform applied to σ_i^2 . So it is scale invariant. In Theil's U-statistic the zero value gives a perfect fit.

4.2.2. Asymmetric Loss Functions

It is a logical assumption that investors will have a different reaction towards negative changes in prices than positive changes. In equity markets this means there is a tendency that the unfortunate news and resulting price falls generate more volatility than the price rise of the same magnitude (Brooks 2002: 438). Investors thus tend to react on a negative rise on uncertainty with greater sensitivity. This phenomenon, known as asset asymmetry, is inverted in the commodities, such as the crude oil is. So the price increase in commodities is a negative incident, where in the share prices the price increase is positive event. This fact was introduced with the other well known and documented stylized facts of volatility in the second main chapter of this thesis. The test on this study is organized so that the asset asymmetry appear the same way in crude oil return series as it would in the equity shares returns.

To address the need of asymmetric metrics in forecast evaluation, some solutions have been suggested in the literature. The one used in this paper is the logarithmic error (LE) introduced by Pagan and Schwert (1990). It is a loss function that penalises volatility forecasts asymmetrically:

$$(21). \quad LE = \frac{1}{T} \sum_{i=1}^T [\ln(\sigma_i^2) - \ln(\hat{\sigma}_i^2)]^2 .$$

Bollerslev and Ghysels (1996) suggested their heteroskedasticity adjusted mean square error (HMSE) statistic as follows:

$$(22.) \quad HMSE = \frac{1}{T} \sum_{i=1}^T \left(\frac{\sigma_i^2}{\hat{\sigma}_i^2} - 1 \right)^2 .$$

In equation (4.7) the \ln is a natural logarithm, $\hat{\sigma}_i^2$ is forecast of realized volatility σ_i^2 , and T denotes number of periods.

Brailsford et al. (1996) suggested mean mixed error statistics (MME). The statistics is the sum of two error statistics modules: MME(U) and MME(O). The first penalises more under predictions and the latter logically penalises over predictions. In the equations, the notation T_O over sigma means over prediction and T_U over sigma under prediction, otherwise notation follows previous equations:

$$(23). \quad MME(U) = \frac{1}{T} \left[\sum_{i=1}^{T_O} |\hat{\sigma}_i^2 - \sigma_i^2| + \sum_{i=1}^{T_U} \sqrt{|\hat{\sigma}_i^2 - \sigma_i^2|} \right],$$

$$(24). \quad MME(O) = \frac{1}{T} \left[\sum_{i=1}^{T_O} \sqrt{|\hat{\sigma}_i^2 - \sigma_i^2|} + \sum_{i=1}^{T_U} |\hat{\sigma}_i^2 - \sigma_i^2| \right].$$

5. EMPIRICAL TESTS AND FINDINGS

In this chapter the empirical tests of this thesis are presented. The last details on the empirical test setup are revealed in the following subchapter. After that follows the remaining details of the test arrangement for volatility models presented under scrutiny. Then, after these last details, is the time for results and the forecasting capability evaluation.

The hypotheses that will be tested in this study are following:

5. More complex model yields more accurate forecasts than simpler one.
6. Asymmetric volatility model results more accurate forecasts than the symmetric model.
7. More volatile period results in inferior volatility forecasts.
8. Volatility forecasting capability decreases with longer horizon.

Volatility modelling is well studied field in contemporary finance and there is several alternative forecast evaluation statistics. Unfortunately in literature there does not seem to be any clear consensus over which one is the best error statistic. To overcome this, several alternative forecast evaluation measures are reported. This also seems to be normal procedure in literature. (Madrakis 1993; Armstrong *et al.* 1992; Brailsford *et al.* 1996).

5.1. Empirical Data

The empirical samples are taken from the daily return series encompassing a period from January 1990 to October 2005. This period as a whole contains several crude oil price moving large scale events making it interesting to contemplate. During this data period the price level for the barrel has more than tripled. The real connection between uncertainty and price development can be seen for example in the early part of the data, from 17th January 1991. Then the Coalition of forces started their operation in order to liberate occupied Kuwait from the Iraqi's. At the markets this was seen as a major change in uncertainty level of an oil supply. During the two following days price levels plunged, hence the negative returns for these days were -22,5% and -16,8% respectively. On those data points culminates the highest return changes in this data set. The

same type of pricing behaviour recurs, but on a smaller scale, with events like a prolonged conflict following the second Gulf War or the exceptionally severe hurricane season in the Gulf of Mexico in 2005. The latter influenced heavily to biggest crude oil consuming nation's oil drilling, refining and storage activities in the region. These market reactions also show the nature of crude oil as commodity priced by its supply, demand and inventory. These price shocks also raise the reason for using shorter estimation periods. This can also be seen from a graphical illustration of prices. The complete Brent Crude Oil Index prices graph can be seen from chapter 4.1 in figure 2.

The whole data set is divided into two diverse estimation periods. The first period is the rather smooth mid 1990's, the latter representing higher level of uncertainty during the early part of the following decade. The idea in the first data set is to test volatility models forecasting performance after normal estimation period. Then the second test is conducted using the same models and estimating parameters on the same length period but during more turbulent times. This should reveal some interesting characteristics on different models capacity to model underlying series under diverse conditions.

As it is industry standard in the financial econometric literature, price data is converted into log-return time series. Diverging from the "normal" asset returns, the returns are inverted due to nature of commodities. The well documented asset price asymmetry is also found in commodities too, but in the opposite way compared to stock market and other "normal" investment goods. The rising price in commodities is the same signal than plunging price at the stock market. Thus to give a better performance possibilities to EGARCH and TARARCH models that are designed to have asymmetric properties, the series is inverted. The log-returns are calculated in the following way

$$(25). \quad r_t = \ln\left(\frac{p_t}{p_{t-1}}\right).$$

Where p_t is price at time t and r_t is the return from day $t-1$ to day t . The complete series is then multiplied with -1 to invert it. The reason for using log-returns is two fold. First, the log-returns can be interpreted as continuously compounded returns and secondly this leads to a time-additive property which is needed in this work. (Brooks 2002: 6 – 8).

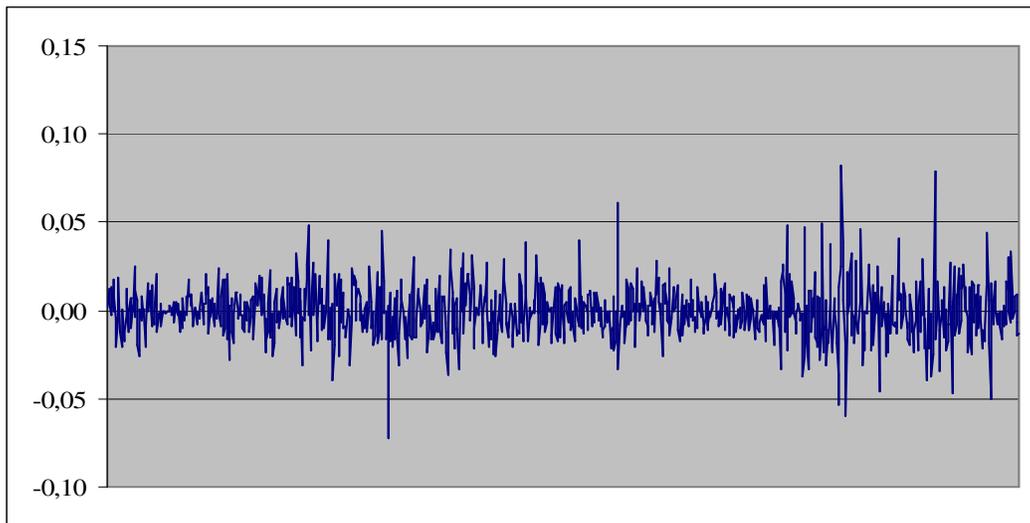


Figure 3. The inverted return series, 4th January 1993 to 31st January 1997.

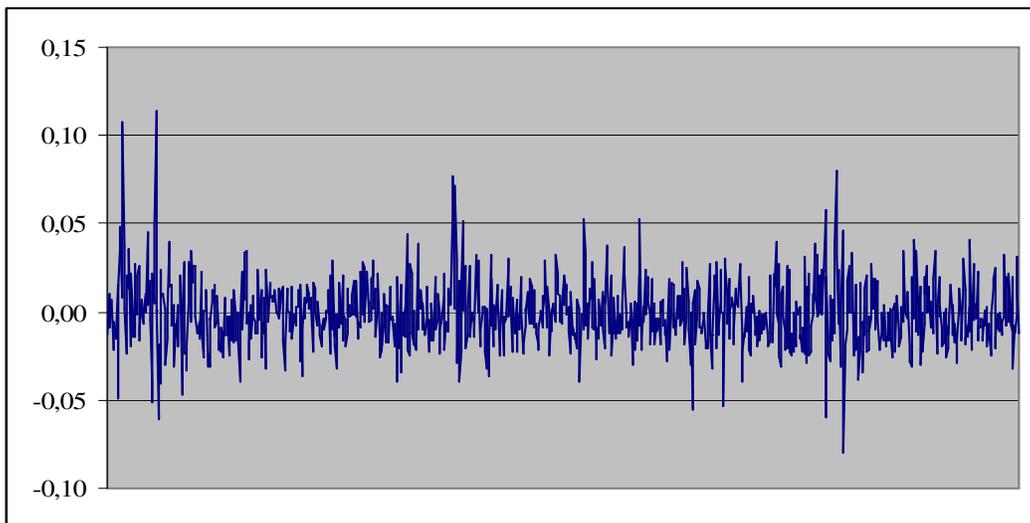


Figure 4. The inverted return series, 30th August 2001 to 28th September 2005.

As the daily Log-return series covered 4111 observations. From this data set, two different periods was extracted. The first period was from the 4th January 1993 to 31st January 1997 or the observations 784 – 1848. This period was used to estimate volatility models tested. The second period covers the observations

3042 – 4106, starting 30th August 2001 and lasting until 28th September 2005. The first period was selected to be more tranquil than the second. This can be easily observed by visually comparing the figures 3 and 4. The both figures have the same scale on their Y-axis. Thus, it is clear that the second period is more volatile through the whole estimation period. It has more frequent clustering, more frequent mean reversion and negative fluctuations tend to go deeper.

The realized volatility is simply computed as the sum of squared daily Brent crude log-returns spread either over one, three or five day forecasting horizon.

$$(26). \quad \sigma_R = \sqrt{\sum_{i=1}^n r_{t+i}^2},$$

where the r is the daily return. The σ_R is realized volatility over the forecast horizon. Until any testing is done all estimates are annualized into the following form:

$$(27). \quad \sigma_A = \sigma_t \sqrt{\frac{252}{T}}.$$

There σ_A is the annualized volatility estimate. 252 is the number of trading days in the year, the nonannualized volatility estimate is σ_t , and the number of days in the forecasting horizon is denoted by T .

The statistical properties for the two time series at hand are represented in figures 3 and 4 below. For the first estimation period the daily inverted returns average at -0,02% and they vary between 8,25% and -7,24%. The second period has the average at -0,08% and varies between 11,35% to -8,02%.

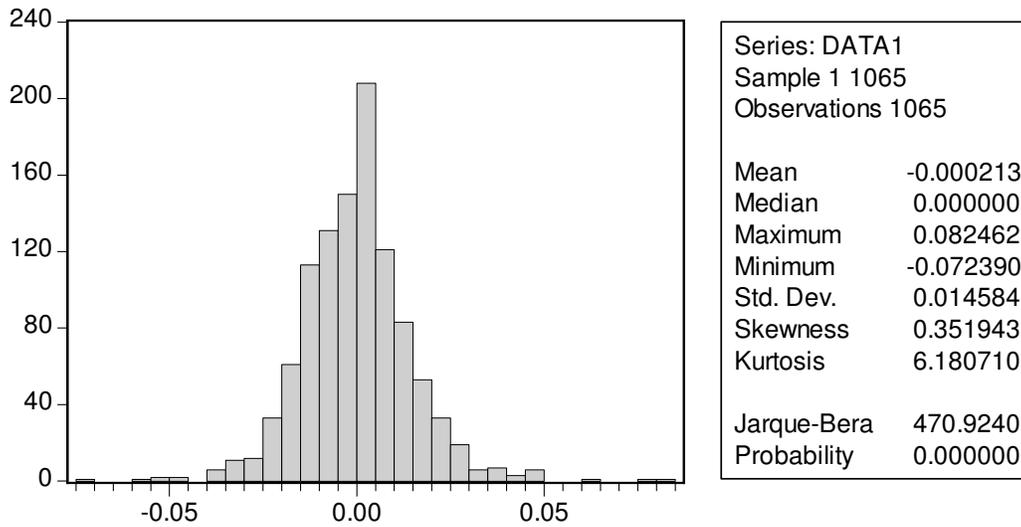


Figure 5. Descriptive statistics for inverted Brent Crude Oil Index returns from 4th January 1993 through 31st January 1997.

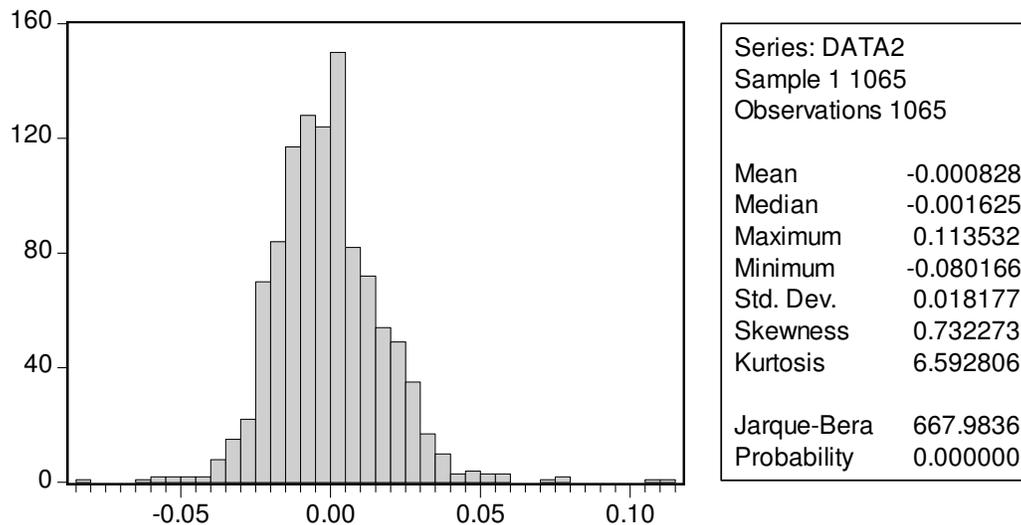


Figure 6. Descriptive statistics for inverted Brent Crude Oil Index returns from 30th August 2001 to 28th September 2005.

The two return series were tested for departures from normality with the Jarque-Bera (1980) test. This test uses the property of a normally distributed

random variable, that the entire distribution is characterized by the first two moments. These naturally are the mean and the variance. It measures departures from normality by the standardized third and fourth moments. They are distributions of skewness and kurtosis. Normal distribution is not skewed and has kurtosis coefficient 3. Large numbers for the Jarque-Bera (1980) statistic will flag significant departures of normality. Both of the tested series indicated in clear numbers that they are not normally distributed. This is normal in financial time series data. (Mills 2000: 223 – 224; Brooks 2002: 179 – 180).

Table 1. The augmented Dickey-Fuller test for first return series.

Augmented Dickey-Fuller Test		t-Statistic	Prob.
Dependent Variable: D(DATA1)			
Augmented Dickey-Fuller test statistic		<u>-23.13613</u>	<u>0.0000</u>
Test critical values:	1% level	-3.436284	
	5% level	-2.864048	
	10% level	-2.568157	

Table 2. The augmented Dickey-Fuller test for second return series.

Augmented Dickey-Fuller Test		t-Statistic	Prob.
Dependent Variable: D(DATA2)			
Augmented Dickey-Fuller test statistic		<u>-26.37708</u>	<u>0.0000</u>
Test critical values:	1% level	-3.436278	
	5% level	-2.864046	
	10% level	-2.568156	

The both daily return data periods were tested for possible unit roots by using the augmented Dickey-Fuller (1979) (later ADF) test. The test statistics are above in tables 1 and 2. The null hypothesis of the ADF test is that tested time series have unit root. The ADF statistic for the first sample was -23.136, thus having significantly lower value than the 1% critical value, -3,436. The second ADF statistic was -26,377 and the critical value was at -3,436, hence the null hypothesis was rejected again. Neither of these return series fulfilled the null hypothesis of the ADF test and had a unit root. Thus both series are stationary.

The Engel's (1982: 1002) ARCH-test for 5 lags was also conducted to returns data. The test results are reported in table 3. The test finds if there is an autocorrelation in the squared residuals. During the both estimation periods of return data, the F-statistic and the LM-statistic suggest presence of the ARCH-effect in return series.

Table 3. Engle ARCH-tests for returns. Both estimation periods as own sample.

ARCH Test:			
<u>Sample 1</u>			
F-statistic	9.641211	Prob. F(5,1054)	0.000000
Obs*R-squared	46.36013	Prob. Chi-Square(5)	0.000000
<u>Sample 2</u>			
F-statistic	10.96185	Prob. F(5,1054)	0.000000
Obs*R-squared	52.39659	Prob. Chi-Square(5)	0.000000

5.2. Volatility Model Parameter Estimates

All volatility model parameters are estimated using the EViews 5 -software. Also the volatility forecasts are produced with this software. The basic assumption in estimation have been that the error distribution is Gaussian and the optimization algorithm is one by Berndt, Hall, Hall, & Hausman (1974) or BHHH as it is widely known in the literature.

When using the EViews 5 for forecasting from estimated GARCH -type specification, there is a choice to use either dynamic or static forecasting method. The Dynamic method calculates a dynamic, multi-step forecasts starting from the first period of the forecast sample. Whilst the static method calculates a sequence of one-step forward forecasts, using the realized, rather than forecasted values, thus all the time updating the process. In the dynamic forecasting, previously forecasted values for the lagged dependent variables are used in forming forecasts of the current value. In this study, the dynamic method is used to produce actual multi-step forecasts. This way there is possible to observe the exact forecasting performance of the models in a given time horizon. Otherwise by using the EViews static method, the metric for the volatility models' predicting power several periods ahead would be lost, the

static forecasts could be interpreted as an ever renewing one day a head forecasts.

In order to conduct a forecasting test on a different GARCH -type volatility models, the first step is to estimate model parameters to each different volatility model. In this study this is done for two estimation periods, thus generating six sets of parameters for the three different models. These parameter estimates are chronologically discussed from the simplest model to the most complex one.

In all three models the (p, q) parameter were set to 1. The decision was based on Akgiray's (1989) founding that in the class of GARCH processes for market volatility, The GARCH $(1,1)$ specification provides the best fit using a likelihood ratio test. Naturally, if the GARCH is preset to $(1,1)$, the other models follow to get a comparable findings.

The first data period for volatility model parameter estimation starts on 4th of January 1993 and goes on until 31st January 1997. It holds total of 1065 return observations which is, by coincidence, the exactly same amount of observations as it is in the later estimation period. The latter period begins 30th August 2001 and stretches until 28th September 2005. The model parameters are discussed starting with the GARCH, then the TARARCH and then ending up to the EGARCH.

Table 4. The GARCH model parameter estimates on first period.

GARCH	Parameter	Value	Prob.
<i>Sample 1</i>	C(1): ω	0,000003	0.0024
<i>observations</i>	C(2): α	0,076971	0.0000
<i>784 - 1848</i>	C(3): β	0,910362	0.0000

Table 5. The GARCH model parameter estimates on second period.

GARCH	Parameter	Value	Prob.
<i>Sample 2</i>	C(1): ω	0,000035	0.0012
<i>observations</i>	C(2): α	0,113273	0.0000
<i>3042 - 4106</i>	C(3): β	0,781250	0.0000

The GARCH model for the first estimation period has as constant parameter 0,000003, the ARCH parameter 0,076971 and the GARCH parameter is 0,910362. From the later period, the constant is 0,000034, the ARCH-term is 0,113273 and the lag term is 0,781250. This clearly shows that later period is a lot more volatile since the short-term information is taken into account more heavily. The first estimation period yielded a sum of alpha and beta to set 0,987 level. This should be compared to the second period, where the corresponding sum was only 0,895 indicating less modelling power from GARCH model to the latter period.

Table 6. The TARARCH model parameter estimates on first period.

TARARCH	Parameter	Value	Prob.
<i>Sample 1</i>	C(1): ω	0,000003	0.0067
<i>observations</i>	C(2): α	0,047596	0.0002
<i>784 - 1848</i>	C(3): γ	0,064325	0.0001
	C(4): β	0,908763	0.0000

Table 7. The TARARCH model parameter estimates on second period.

TARARCH	Parameter	Value	Prob.
<i>Sample 2</i>	C(1): ω	0,000079	0.0000
<i>observations</i>	C(2): α	0,231153	0.0000
<i>3042 - 4106</i>	C(3): γ	-0,203786	0.0000
	C(4): β	0,627345	0.0000

The TARARCH model has in the first period the following values. The constant is 0,000003, the ARCH-term is 0,047596, the asymmetry capturing term with dummy term acting as the asymmetry switch is 0,064325 and the GARCH-term is 0,908763. In the latter estimation period, the same terms in the same order are: 0,000078, 0,231153, -0,020378 and 0,627345. The same pattern in the normal GARCH estimates is also evident here. The later period is clearly more volatile, thus the long-term memory retaining GARCH-term is also here significantly lower than during the first estimation period.

Table 8. The EGARCH model parameter estimates on first period.

EGARCH	Parameter	Value	Prob.
<i>Sample 1</i>	C(1): ω	-0,305887	0.0000
<i>observations</i>	C(2): α	0,150989	0.0000
<i>784 - 1848</i>	C(3): γ	-0,049182	0.0000
	C(4): β	0,977619	0.0000

Table 9. The EGARCH model parameter estimates on second period.

EGARCH	Parameter	Value	Prob.
<i>Sample 2</i>	C(1): ω	-1,954326	0.0000
<i>observations</i>	C(2): α	0,215267	0.0000
<i>3042 - 4106</i>	C(3): γ	0,132566	0.0000
	C(4): β	0,777942	0.0000

When the EGARCH estimation results are analyzed, the same pattern arises as it was with the other two models. The first period terms are in first: -0,305887, 0,150989, -0,049182 and the long-term volatility memory is 0,977619. The second estimation period yielded following parameter values: -1,954326, 0,215267, 0,132566 and 0,777942. As in the other models the longer term memory has lesser effect in the second period, giving away some incriminating evidence for the more volatile period.

After computing the estimates, it is time to make forecasts. Then when these forecasts are compared to actual values, it is possible to see if any forecasting method outperforms others.

5.3. Empirical Results and Forecast Evaluation

The forecast accuracy is unveiled in the following part by using various statistics to evaluate the forecasting capability of each volatility model tested. The evaluation statistics are computed for 1-, 3-, and 5-day volatility forecasts. The statistics include both symmetric and asymmetric forecast evaluation methods, two on latter group and three on symmetric statistics. On the other hand, in the symmetric set there is both AMAPE and MAPE reported, so one could argue that there is actually used only two statistics on symmetric forecast evaluation methods. The reason why both of them are reported is curiosity to see how big of a difference is there between these two statistics.

The first estimation period between 4th January 1993 and 31st January 1997 was used to forecast the results seen in tables 10 to 12.

Table 10. The first estimation and forecasting period, 1-day forecasts.

1-day forecast	RMSE	MAPE	AMAPE	LE	HMSE
GARCH	0,0001449	390,75 %	132,29 %	2,53	0,6340
TARCH	0,0001302	351,00 %	127,40 %	2,27	0,6057
EGARCH	0,0001417	382,02 %	131,27 %	2,47	0,6281

Table 11. The first estimation and forecasting period, 3-day forecasts.

3-day forecast	RMSE	MAPE	AMAPE	LE	HMSE
GARCH	0,0000022	0,23 %	0,23 %	0,000016	0,000016
TARCH	0,0000270	2,82 %	2,95 %	0,002610	0,002854
EGARCH	0,0000074	0,77 %	0,78 %	0,000181	0,000185

Table 12. The first estimation and forecasting period, 5-day forecasts.

5-day forecast	RMSE	MAPE	AMAPE	LE	HMSE
GARCH	0,00005793	2,47 %	2,63 %	0,0035	0,0040
TARCH	0,00008920	3,80 %	4,19 %	0,0089	0,0110
EGARCH	0,00006408	2,73 %	2,93 %	0,0043	0,0050

When making the 1-day forecasts from this sample, the TARCH model dominated. However, the results for 3- and 5-day forecasts did not support this finding. When the forecasting horizon got longer, the GARCH model seems to yield smallest forecast error.

The second period was estimated from the data set including observations from 30th August 2001 to 28th September 2005. Again forecasts for the 1-, 3-, and 5-day forecasting periods were calculated. The forecast evaluation statistics are in the tables 13 to 15.

Table 13. The second estimation and forecasting period, 1-day forecasts.

1-day forecast	RMSE	MAPE	AMAPE	LE	HMSE
GARCH	0,0003078	12340,64 %	196,81 %	23,27	0,9840
TARCH	0,0003280	13149,44 %	197,00 %	23,88	0,9850
EGARCH	0,0003483	13965,86 %	197,18 %	24,47	0,9858

Table 14. The second estimation and forecasting period, 3-day forecasts.

3-day forecast	RMSE	MAPE	AMAPE	LE	HMSE
GARCH	0,0004805	266,65 %	53,33 %	1,61	0,2634
TARCH	0,0005093	282,62 %	53,94 %	1,69	0,2667
EGARCH	0,0005386	298,89 %	54,51 %	1,76	0,2698

Table 15. The second estimation and forecasting period, 5–day forecasts.

5-day forecast	RMSE	MAPE	AMAPE	LE	HMSE
GARCH	0,0003009	15,02 %	10,92 %	0,0628	0,0368
TARCH	0,0003318	16,57 %	11,71 %	0,0728	0,0410
EGARCH	0,0003640	18,18 %	12,50 %	0,0836	0,0453

Now the basic GARCH –model dominates in every forecast category. The TARCH –model gave second best results leaving the EGARCH –model most inaccurate in this test arrangement.

In the both test arrangements it seems to be clear that the GARCH –model yields the most accurate forecasts. This came rather surprisingly, since the two asymmetric models (the TARCH and the EGARCH) are designed to capture properties of the time series more accurately and therefore should outperform the basic GARCH –model.

As a by product came observation that the two forecast evaluation statistics, MAPE and AMAPE, give increasingly different results as the forecasting error increases. By these very limited observations, the adjusted MAPE seems to be a better statistic if there can be expected some rather large forecast errors. This study has unfortunately too small sample to make any further deduction on this matter.

The first hypothesis was that the more complex model yields more accurate forecasts than simpler one. Thus the EGARCH should produce the most accurate forecasts. Then after that should the TARCH be better forecast yielding model before the basic GARCH –model. Complexity in this context should be understood as accumulation of more parameters to the model. Only in the first forecast sample and only with the 1–day forecast, the TARCH was most accurate model. Thus second most complex model had its moment in there. Even then the most elaborate model, the EGARCH was left to second in forecast accuracy. Elsewhere on the other forecast lengths (3–day and 5–day) the GARCH –model dominates in first test. In the second set of forecasts, the GARCH –model is superior to the other two models tested. With this found

evidence, there no alternative but to reject the first hypothesis, the simplest model was the most accurate one.

The second hypothesis was that the asymmetric volatility model results more accurate forecasts than the symmetric model. This hypothesis was included to see if models appear in different order in comparison to the first hypothesis. The second hypothesis was as less fortunate as the first one. The hypothesis should also be discarded, since the symmetric GARCH -model dominated in the forecasting accuracy during the both tests.

The third hypothesis gets sound backing from the findings. This clearly shows when comparing forecast errors from less volatile first testing period, to second, more volatile period. All the forecast error statistics yield significantly higher error levels on the more volatile test period. The hypothesis was confirmed by the findings.

The last hypothesis stated that the volatility forecasting capability is linked to forecasting horizon, namely the models lose the capability to produce accurate forecasts over time. This hypothesis yielded rather surprising findings. Naturally all these three models tested will lose their accuracy over time since they all are path dependent history based models and the time series to be forecasted absorbs all the time new information. Thus it was surprising to find, when moving from the 1-day forecast to the 5-day forecasts, the forecasting accuracy was getting better. This can be logically explained by nature of these forecasting models. The forecast is only about magnitude of following volatilities, in real world the return levels can and will change rather randomly. When horizon goes to 5-days, the amount of observations is sufficient enough to average forecasts, hence yielding more accurate forecasts. To test this fourth hypothesis more thoroughly it, the forecast horizon should be longer. By the findings gotten for these two sets of forecasting accuracy tests, the hypothesis is discarded; the forecast error levels get smaller as it is gone from the 1-day forecasts to the 5-day forecasts.

There can be some reasons why the results were bit mixed and most of the hypotheses could not be confirmed. Mainly the reason could lie in rather limited test sample. Since only one business week was forecasted ahead from estimation period in both cases, the sample size can be too limited. It is also

widely known in literature that the GARCH -effect does fade away as the frequency of observations gets lower (see Anderssen *et al.* 1997; Engle 2000). The forecasts in this light might get more accurate if the frequency in data would be higher.

6. CONCLUSIONS

The purpose of this study is to find out the forecasting performance of certain class of econometrical models applied to crude oil return data. These three ARCH -family models the GARCH, the TARARCH, and the EGARCH were tested to elicit forecasting capability embedded in them. The latter two of the models were asymmetric by their nature, hence giving further potential for fitting to the oil series properties. Rather surprisingly the symmetrical and the simplest GARCH -type specification got best of the lot. The Brent Crude Oil index data covered daily closing prices from January 1990 to October 2005. From this data set it was extracted two separate periods for estimating the volatility model parameters. Then the return series was inverted to address data asymmetry problem. The inverted return series sets were tested and then subjected to the volatility model estimation. After the both periods, a set of forecasts is generated. The forecasting length is set to 1, 3, and 5-days. Forecast evaluation methods are applied in order to find the smallest forecasting errors.

Four hypotheses were formed in this study based on the findings in earlier studies. The first hypothesis suggests that the more complex model should generate most accurate forecasts. In this context, the growing complexity is understood as accumulation of more parameters in model. The forecast evaluation statistics do not support such assumption and the hypothesis is rejected.

The second hypothesis inspected if the asymmetric volatility model results more accurate forecasts than the symmetric model. The hypothesis should also be discarded, as the error statistics do not support this hypothesis. The third hypothesis was more successful. It stated that a more volatile period results inferior volatility forecasts. Comparing the error statistics from the less volatile first test period to the more volatile second test period confirms the hypothesis. The less volatile period assists these volatility models give more accurate forecasts.

The fourth and the final hypothesis stated that the volatility forecasting capability is linked to forecasting horizon, namely the models lose their capability to produce accurate forecasts over time. This hypothesis yielded rather surprising findings. Naturally all these three models tested will lose their

accuracy over time since they all are path dependent history based models and the time series to be forecasted absorbs all the time new information. Thus it was surprising to find, when moving from the 1-day forecast to the 5-day forecasts, the forecasting accuracy was actually getting better.

REFERENCES

- Ahlstedt, M. (1998). *Analysis of Financial Risks in a GARCH Framework*. 1st ed. Helsinki: Bank of Finland Studies E:11. 181 p.
- Akgiray, V. (1989). Conditional Heteroskedasticity in time series of stock returns: evidence and forecast. *Journal of Business* 62. 55 – 80.
- Anderssen, T & T. Bollerslev (1997). Heterogeneous Information Arrivals and Return Volatility Dynamics: Uncovering the Long-run in High Frequency Returns. *Journal of Finance*. 52, 975 – 1005.
- Andersen, T. & T. Bollerslev (1998a). Deutsche Mark-Dollar volatility: intraday volatility patterns, macroeconomic announcements, and longer run dependencies. *Journal of Finance* 53. 219 – 265.
- Andersen, T. & T. Bollerslev (1998b). Answering the sceptics: yes, standard volatility models do provide accurate forecasts. *International Economic Review* 39. 885 – 905.
- Andersen, T., Bollerslev T., Diebold F. & P. Labys (2001). The distribution of exchange rate volatility. *Journal of American Statistical Association* 96. 42 – 55.
- Alexander, C. (2001). *Market Models. A Guide to Financial Data Analysis*. 1st ed. Chichester: John Wiley & Sons Ltd. 494 p.
- Armstrong J. & F. Collopy (1992). Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting* 8. 69 – 80.
- Baillie, R., Bollerslev, T. & H. Mikkelsen (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 74. 3 – 30.

- Berndt, E., Hall, B., Hall, R. & J. Hausman (1974). Estimation and inference in nonlinear structural models. *Annals of Economic and Social Measurement* 3. 653 – 665.
- Black, F. (1976). The Pricing of Commodity Options. *The Journal of Financial Economics*. 3, 167 – 179.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31. 307 – 327.
- Bollerslev, T., Chou, R. & K. Kroner (1992). ARCH modelling in finance. *Journal of Econometrics* 52. 5 – 59.
- Bollerslev, T., Engle, R. & D. Nelson (1994). ARCH models. In: *Handbook of Econometrics vol IV*, 2959 – 3038. Ed. Engle R. & D. McFadden. Amsterdam: North-Holland.
- Bollerslev, T. & E. Ghysels (1996). Periodic autoregressive conditional heteroskedasticity. *Journal of Business and Economic Statistics* 14. 139 – 157.
- Bollerslev, T. & M. Melvin (1994). Bid-ask spreads and the volatility in the foreign exchange market: an empirical analysis. *Journal of International Economics* 36. 355 – 372.
- Bollerslev, T. & J. Wooldridge (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Review* 11. 143 – 172.
- Brailsford, T. & R. Faff (1996). An evaluation of volatility forecasting techniques. *Journal of Banking and Finance* 20. 419 – 438.
- Brooks, C. (1998). Predicting stock index volatility: Can market volume help? *Journal of Forecasting* 17. 59 – 80.
- Brooks, C. (2002). *Introductory Econometrics for Finance*. 1st ed. Cambridge: Cambridge University Press. 701 p.

- Brenner, R., Harjes, R. & K. Kroner (1996). Another look at models of the short-term interest rate. *Journal of Financial Quantitative Analysis* 31. 85 – 107.
- Chan, K., Karolyi G., Longstaff F. & A. Sanders (1992). An empirical comparison of alternative models of the short-term interest rate. *Journal of Finance* 47. 1209 – 1227.
- Chou, R. (1988). Volatility persistence and stock valuations: some empirical evidence using GARCH. *Journal of Applied Econometrics* 3. 279 – 294.
- Christie, A. (1982). The stochastic behavior of common stock variances: value, leverage and interest rate effects. *Journal of Financial Economics* 10. 407 – 432.
- Cootner, P. (1964). *The random character of stock market prices*. Cambridge: MIT Press.
- Day, T. & C. Lewis (1993). Forecasting futures market volatility. *Journal of Derivatives* 1. 33 – 50.
- Dickey, D. & W. Fuller (1979). Distribution of the estimators for the autoregressive time series with a unit root. *Journal of the American Statistical Association* 74. 427 – 431.
- Dielman, T. (1986). A comparison of forecasts from least absolute value and least squares regression. *Journal of Forecasting* 5. 189 – 195.
- Drost F. & T. Nijman (1993). Temporal aggregation of GARCH processes. *Econometrica* 61. 909 – 927.
- Duan, J-C (1995). The GARCH option pricing model. *Mathematical Finance* 5. 13 – 32.

- Ederington, L. & J. Lee (1995). The Short-run Dynamics of the Price Adjustment to New Information. *Journal of Financial and Quantitative Analysis*. 30, 117 - 134.
- Engle, R. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica* 50. 987 - 1008.
- Engle, R. & T. Bollerslev (1986). Modelling the persistence of conditional variances. *Economic Reviews* 5. 1 - 50.
- Engle, R., Lilien, D., & R. Robins (1987). Estimation of time varying risk premia in the term structure: the ARCH-M model. *Econometrica* 55. 391 - 407.
- Engle, R., Ito, T. & W-L Lin (1990). Meteor showers or heat waves? Heteroskedastic intra-day volatility in the foreign exchange market. *Econometrica* 58. 525 - 542.
- Engle, R. & V. Ng (1993). Measuring and testing the impact of news on volatility. *Journal of Finance* 48. 1749 - 1778.
- Engle, R. (2000). The econometrics of ultra high frequency data. *Econometrica* 68. 1 - 22.
- Engle, R. (2001). GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics. *Journal of Economic Perspectives* 15. 157 - 168.
- Engle R. & A. Patton (2001). What good is a volatility model? *Quantitative Finance* 1. 237 - 245.
- Fama, E. (1965). The behaviour of stock market prices. *Journal of Business* 38. 34 - 105.
- Fama, E. (1970). Efficient capital markets: review of theory and empirical work. *The Journal of Finance*. 25. 383 - 432.

- Fama, E. (1991). Efficient capital market: II. *The Journal of Finance*. 46. 1575 – 1613.
- Franses, P. (2000). *Time series models for business and economic forecasting*. 1st ed. Cambridge: Cambridge University Press. 280 p.
- Geman, H. (2005). *Commodities and commodity derivatives: modelling and pricing for agriculturals, metals, and energy*. 1st ed. Chichester: John Wiley & Sons Ltd. 396 p.
- Glosten, L., Jagannathan, R & D. Runkle (1993). On the relation between expected value and the volatility of the nominal excess returns on stocks. *The Journal of Finance* 48. 1779 – 1801.
- Goodhart, C. & M. O'Hara (1997). High frequency data in financial markets: issues and applications. *Journal of Empirical Finance* 4. 73 – 114.
- Hagerud, G. (1997). Modelling Nordic stock returns with asymmetric GARCH models[online]. *Stockholm School of Economics Working Paper Series in Economics and Finance* 164. Available from World Wide Web: <URL: <http://swopec.hhs.se/hastef/papers/hastef0164.pdf>>
- Hamilton, J. & R. Susmel (1994) Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics* 64. 307 – 333.
- Hansen, B. (1994). Autoregressive conditional density estimation. *International Economic Review* 35. 705 – 730.
- Harvey, C. & A. Siddique (1999). Autoregressive conditional skewness. *Journal of Financial Quantitative Analysis* 34. 465 – 487.
- Higgins, M: & A. Bera (1992). A class of nonlinear ARCH models. *International Economic Review* 33. 137 – 158.
- Hull, J. (2000). *Options, futures & other derivatives*. 4th ed. London: Prentice-Hall International. 698 p.

- Jarque, C. & A. Bera (1980). Efficient tests for normality, homoskedasticity and serial dependence of regression residuals. *Economics Letters* 6. 255 – 259.
- Kaldor, N. (1939). Speculation and economic stability. *Review of Economic Studies* 7. 1 – 27.
- Liljeblom E. & M. Stenius (1993). Macroeconomic volatility and stock market volatility; empirical evidence on Finnish data. *Meddelanden från Svenska handelshögskolan*. 278.
- Liljeblom E. & M. Stenius (1995). Macroeconomic volatility and stock market volatility; empirical evidence on Swedish data. *Meddelanden från Svenska handelshögskolan*. 296.
- Lumsdaine, R. (1995). Finite sample properties of the maximum likelihood estimation in GARCH (1,1) and IGARCH (1,1) models: A Monte Carlo investigation. *Journal of Business Economics and Statistics* 13. 1 – 10.
- Madrakis, S. (1993). Accuracy measures: Theoretical and practical concerns. *International Journal of Forecasting* 9. 527 – 529.
- Markowitz, H. (1952). Portfolio selection. *Journal of Finance* 7. 77 – 91.
- McMillan, D., Speigh A. & O. Gwilym (2000). Forecasting UK stock market volatility. *Applied Financial Economics* 10. 435 – 448.
- Miller M., Muthuswamy & R. Whaley (1994). Mean reversion of Standard & Poor's 500 index basis changes: Arbitrage-Induced or Statistical Illusion? *Journal of Finance*. 49, 479 – 513.
- Mills, T. (2000). *The econometric modelling of the financial time series*. 2nd ed. Cambridge: Cambridge University Press. 372 p.
- Nelson, D (1990). ARCH models as diffusion approximations. *Journal of Econometrics* 45. 7 – 38.

- Nelson, D. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*. 59, 347 - 370.
- Nikkinen, J. & P. Sahlström (2004). Scheduled Domestic and U.S. Macroeconomic News and Stock Valuation in Europe. *Journal of Multinational Financial Management* 14, 201 - 215.
- Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business* 36. 394 - 419.
- Organization of the Petroleum Exporting Countries (OPEC) (2006). *OPEC Market Indicators October 2006* [online]. Vienna: Austria. Available from World Wide Web: <URL:<http://www.opec.org/home/Market%20Indicators/2006/pdf/MI102006.pdf>>
- Osborne, M. (1959). Brownian motion in the stock market. *Operations Research* 7. 145 - 173.
- Pagan, A. & G. Schwert (1990). Alternative models for conditional stock volatility. *Journal of Econometrics* 45. 267 - 290.
- Pearson, K. & L. Rayleigh (1905). The problem of random walk. *Nature* 72. 294 - 318.
- Pindyck, R. & D. Rubinfeld (1998). *Econometric models and economic forecasts*. 4th ed. New York: McGraw-Hill. 672 p.
- Pindyck, R. (2001). *Volatility and commodity price dynamics* [online]. MIT Centre for Energy and Environmental Policy Research. Working Paper. Available from World Wide Web: <URL: <http://web.mit.edu/ceepr/www/2001-007.pdf>>.
- Poon, S.-H. & C. Granger (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature* 41. 478 - 539.

- Schwartz, E. (1997). The stochastic behavior of commodity prices: implications for valuation and hedging. *Journal of Finance* 52. 923 – 973.
- Schwert, G. (1989). Why does stock market volatility change over time? *Journal of Finance* 44. 1115 – 1153.
- Sharpe, W. (1964). Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance* 19. 425 – 442.
- Taylor, S. (1994). Modelling stochastic volatility: A review and comparative study. *Mathematical Finance* 4. 183 – 204.
- Teräsvirta, T. (1996). *Two stylized facts and the GARCH (1,1) model*. 1st ed. Stockholm: Stockholm School of Economics Working Paper Series in Economics and Finance. 27p.
- Theil, H. (1966). *Applied Economic Forecasting*. 1st ed. Amsterdam: North-Holland. 474p.
- Tobin, J. (1958). Liquidity preference as behaviour toward risk. *Review of Economic Studies* 25. 65 – 86.
- Wang, P. (2003). *Financial Econometrics: Methods and models*. 1st ed. London: Routledge. 178 p.
- Williams, J. & B. Wright (1991). *Storage and Commodity Markets*. 1st ed. Cambridge: Cambridge University Press. 518 p.
- Wilmott, P. (2001). *Paul Wilmott introduces quantitative finance*. 1st ed. Chichester: John Wiley & Sons. 521 p.
- Working, H. (1949). The theory of price of storage. *American Economic Review* 39. 1254 – 1262.
- Yu J. (2002). Forecasting volatility in New Zealand stock market. *Applied Financial Economics* 12. 193 – 202.

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