Turun yliopisto
University of Turku

ABSTRACT

X	Master´s thesis
	Licentiate´s thesis
	Doctor´s thesis

Subject	Economics	Date	17.1.2014
Author(s)	Lauri Laaksonen	Student number	415459
		Number of pages	64 p.
Title	The Forecast Performance of Model-Free Implied Volatility: Evidence from DAX Index Options		
Supervisor(s)	Prof. Heikki Kauppi		

Abstract

Volatility has a central role in various theoretical and practical applications in financial markets. These include the applications related to portfolio theory, derivatives pricing and financial risk management. Both theoretical and practical applications require good estimates and forecasts for the asset return volatility. The goal of this study is to examine the forecast performance of one of the more recent volatility measures, model-free implied volatility.

Model-free implied volatility is extracted from the prices in the option markets, and it aims to provide an unbiased estimate for the market's expectation on the future level of volatility. Since it is extracted from the option prices, model-free implied volatility should contain all the relevant information that the market participants have. Moreover, model-free implied volatility requires less restrictive assumptions than the commonly used Black-Scholes implied volatility, which means that it should be less biased estimate for the market's expectations. Therefore, it should also be a better forecast for the future volatility.

The forecast performance of model-free implied volatility is evaluated by comparing it to the forecast performance of Black-Scholes implied volatility and GARCH(1,1) forecast. Weekly forecasts for six years period were calculated for the forecasted variable, German stock market index DAX. The data consisted of price observations for DAX index options. The forecast performance was measured using econometric methods, which aimed to capture the biasedness, accuracy and the information content of the forecasts.

The results of the study suggest that the forecast performance of model-free implied volatility is superior to forecast performance of GARCH(1,1) forecast. However, the results also suggest that the forecast performance of model-free implied volatility is not as good as the forecast performance of Black-Scholes implied volatility, which is against the hypotheses based on theory. The results of this study are consistent with the majority of prior research on the subject.

Key words	volatility, financial markets, financial derivatives, econometrics
Further information	

Turun yliopisto
University of Turku

TIIVISTELMÄ

X	Pro gradu -tutkielma
	Lisensiaatintutkielma
	Väitöskirja

Oppiaine	Kansantaloustiede	Päivämäärä	17.1.2014
Tekijä(t)	Lauri Laaksonen	Matrikkelinumero	415459
		Sivumäärä	64 s.
Otsikko	Osakemarkkinoiden volatiliteetin ennustaminen mallivapaan implisiittisen volatiliteetin avulla		
Ohjaaja(t)	Prof. Heikki Kauppi		

Tiivistelmä

Arvopaperin tuoton vaihtelua mittaavalla volatiliteetilla on keskeinen asema monissa rahoitusmarkkinoiden teoreettisissa ja käytännön sovelluksissa. Nämä sovellukset liittyvät muun muassa portfolioteoriaan, johdannaissopimusten hinnoitteluun sekä rahoitusriskien hallintaan. Monet näistä sovelluksista edelllyttävät ennusteiden laatimista tulevalle volatiliteetin tasolle. Tämän tutkimuksen tavoitteena oli selvittää kuinka hyvin ns. mallivapaa implisiittinen volatiliteetti sopii osakemarkkinoiden volatiliteetin ennustamiseen.

Mallivapaa implisiittinen volatiliteetti on optioiden hinnoista johdettu volatiliteettiennuste, joka pyrkii mahdollisimman virheettömästi kuvaamaan markkinoiden odotuksia tulevasta volatiliteetin tasosta. Tästä syystä sen tulisi pitää sisällään kaikki se oleellinen informaatio, mikä optiomarkkinoiden osapuolilla on. Mallivapaan implisiittisen volatiliteetin tulisi kuvata markkinoiden odotuksia tulevasta volatiliteetin tasosta paremmin kuin usein käytetyn Black-Scholes implisiittisen volatiliteetin, koska sen johtaminen perustuu vähemmän rajoittaviin oletuksiin. Näin ollen voisi myös olettaa, että se ennustaa tulevaa volatiliteettia paremmin.

Mallivapaan implisiittisen volatiliteetin ennustekykyä arvioitiin vertaamalla sen tuottamia ennusteita Black-Scholes implisiittisen volatiliteetin ja GARCH(1,1)-mallin tuottamiin ennusteisiin. Ennustettava muuttuja oli saksalainen osakemarkkinaindeksi DAX, jolle laadittiin viikottaiset ennusteet kuuden vuoden aikajaksolle. Tutkielman aineisto koostui pääasiassa DAX-indeksioptioiden hinnoista. Ennustekykyä mitattiin soveltamalla ekonometrisia menetelmiä, jotka pyrkivät mittamaan ennusteiden sisältämää harhaa, tarkkuutta ja informaatiosisältöä.

Tutkimuksen tulosten perusteella mallivapaan implisiittisen volatiliteetin ennustekyky on tilastollisesti merkitsevästi parempi kuin GARCH(1,1)-mallin ennusteiden. Toisaalta mallivapaan implisiittisen volatiliteetin ennustekyky oli tutkimuksessa heikompi kuin Black-Scholes implisiittisen volatiliteetin ennustekyky, vastoin teoriaan perustuvia hypoteeseja. Valtaosa aiheeseen liittyvistä aiemmista tutkimukista on päätynyt samaan lopputulemaan, vaikka poikkeaviakin tuloksia löytyy.

Asiasanat	volatiliteetti, rahoitusmarkkinat, johdannaismarkkinat, ekonometria
Muita tietoja	

Turun yliopisto
University of Turku

THE FORECAST PERFORMANCE OF MODEL-FREE IMPLIED VOLATILITY

Evidence from DAX index options

Master´s Thesis
in Economics

Author:
Lauri Laaksonen

Supervisor:
Prof. Heikki Kauppi

17.1.2014
Turku

Turun kauppakorkeakoulu • Turku School of Economics

Table of contents

List of figures

List of tables

1 INTRODUCTION

One of the central ideas of economic thought is that, in properly functioning markets, prices contain valuable information that can be used to make a wide variety of economic decisions. (Rubinstein 1994, 771)

Volatility forecasting is little like predicting whether it will rain: You can be correct in predicting the probability of rain, but still have no rain. (Engle 1993, 72)

1.1 Volatility forecasting in general

Volatility refers to a spread of different outcomes of a random variable. In financial context, the idea of volatility is to measure the typical size of a change in return on specific asset. The measures of volatility are generally used to capture the uncertainty in financial marketplace. Since the variation in financial risk factors is important for numerous theoretical models of financial markets, volatility is one of the most central concepts in modern financial theory. Nevertheless, volatility is not solely a theoretical construct. In addition to its widespread use in theoretical context, volatility has an important role in a large number of practical applications, such as portfolio management, option pricing and risk evaluation. Therefore, both the theoretical and practical applications require good estimates and forecasts for the asset return volatility.

Volatility is often considered synonymous with risk. In times of high volatility, the security values are not dependable, which indicates that the market is not functioning well and there is uncertainty in the marketplace. Thus, high volatility is seen as an undesirable for a rational risk-averse investor. For example, in Markowitz portfolio selection, the financial asset’s return volatility is used as a measure of risk for that specific asset. However, volatility is not completely the same as risk, if risk is solely associated with a negative outcome. As volatility refers to a spread of different outcomes for the return on financial asset, it is strictly speaking a measure of uncertainty instead of a measure of risk.

It is a standard procedure in economics to view the economic time series as realizations of stochastic processes. The volatility of the economic time series is commonly observed to be changing over time and hence the volatility itself can be seen as a stochastic variable. The research on the financial return series has come up with several stylized facts for the volatility of asset returns. First, the volatility tends to cluster, i.e. the large change in asset returns is typically followed by another large change, and a small change in return is typically followed by another small change. Moreover, the volatility time series tend to have a long memory, which means that the autocorrelation in volatility series is typically significant even for the long lags. Second, volatility tends to be negatively correlated with the asset return, and this correlation is observed to be different for the positive and negative returns. The negative returns typically cause a significant increase in volatility whereas the positive returns tend to cause a smaller negative impact in asset's volatility. Furthermore, the empirical research indicates that the possibility of extreme events in the asset prices is far greater than the modern financial theory suggests. In other words, the typically observed return distributions tend to be fat tailed, even though they are usually assumed to be normally distributed. (see, for example Masset 2011, 1)

In financial context, the most common statistical measure for volatility is the sample standard deviation of the asset return. Since standard deviation has the same unit of measure than mean, it can be interpreted as a typical deviation from the mean. Variance, the squared form of the standard deviation, is also used as a measure of volatility, but it is much less stable and does not have as intuitive interpretation as standard deviation. Hence, standard deviation is more common and convenient measure for volatility. However, the use of standard deviation to measure volatility is not always sensible in practical applications. That is because the sample mean is not an accurate estimate for true mean unless the sample period is very long. Inaccurate mean leads to a noisy volatility estimate as the sample standard deviation is calculated in terms of deviations from the sample mean. On the other hand, use of longer sample periods in the estimation of standard deviation can also be problematic. Even though including older data leads to a larger sample and less noisy estimate for mean, it may contain irrelevant information on the current state of volatility, since volatility is not constant over time. (Poon 2005, 1; Figlewski 1997, 82)

The forecasting of financial market volatility has been a widely studied subject for the past few decades. The forecasting ability of the most popular models has been under a plentitude of research and various new models have been developed. Nevertheless, forecasting the future level of a time-varying variable is usually a difficult task. Therefore it is worthwhile to ask, how well the proposed volatility models forecast future volatility. If the volatility forecasting models do not perform well, volatility forecasting might not be sensible at all. However, the forecasting results in the academic literature indicate that the volatility forecasting is indeed sensible, since the forecasts generated by the volatility models correlate closely with the future latent volatility factor (see, for example Andersen & Bollerslev 1998).

Generally speaking, the academic research on volatility forecasting can be divided into two lines: forecasting with models based on financial time series and forecasting with models that use the information in the option (and other derivatives) markets. The first line of research – the time series modeling of stochastic volatility – was initiated by

Robert Engle (1982). Engle designed the autoregressive conditional heteroskedasticity models (ARCH) to capture the clustering volatility observed in the financial time series. Even though the clustering volatility property was well-known before ARCH models, Engle was the first one to find an efficient way of modeling it. The difference between ARCH models and the earlier models is the notion, that even if the unconditional error variance of a model with stochastic variables is constant, its conditional error variance may well change over time. This notion made the joint-estimation of parameters of conditional mean and variance possible. Bollerslev (1986) introduced a natural generalization to ARCH model by allowing the past conditional variances in the current conditional variance equation. The outcome, generalized autoregressive conditional heteroskedasticity (GARCH) model, and its numerous extensions are still widely used in volatility forecasting. In addition to the popular conditional heteroskedasticity models, various other ways of modeling and forecasting volatility by means of historical time series have been proposed.

The second line of volatility forecasting research – the implied volatility measures extracted from the option prices – is originally a by-product of Black-Scholes pricing equation for European options (Black & Scholes 1973). The pricing equation, which has gained significant popularity in both theoretical and practical work, gives a theoretical price for a European option as an outcome of five variables. The central idea of the pricing equation is to find such price for the option that there are no arbitrage opportunities, given that the assumptions of the model are valid. The model is particularly attractive because almost all of its parameters are observable in the marketplace. The only input that is not directly observable is the future return volatility of the underlying asset. One way to estimate it is to use Black-Scholes pricing equation backwards: since the option prices in the market are observable, it is possible to find the level of volatility which is consistent with the current market price. Given that the Black-Scholes pricing equation is the "correct" model for option pricing, this volatility level implied by the option prices, is the market expectation of the underlying asset's volatility over the life of the option. This backward working use of the pricing equation and its output – Black-Scholes implied volatility – is often more useful than the theoretical price of the option. While it is a standard approach to use Black-Scholes implied volatility as a market's expectation for the future volatility, several other ways of extracting the implied volatility from the option prices have been proposed.

The forecasting power of the implied volatility measures is widely considered to be superior to the forecasts generated with time series models. That is because they embody the market participants' expectations of the future volatility. Hence, implied volatility measures are considered to subsume all the information in the time series forecasts (see, for example Jiang & Tian 2005, 1305). Nevertheless, the most common implied volatility measure – Black-Scholes implied volatility – has certain misgivings. To begin with, Black-Scholes model has a number of restrictive assumptions, which are inconsistent with the stylized facts of the financial markets. For example, the model assumes underlying asset's volatility to be constant and its return distribution to be lognormal. Due to these assumptions, the Black-Scholes pricing equation is not likely to be the "correct" model for option pricing, which means that the implied volatility cannot be treated as an unbiased estimate of the market's expectations on the future volatility. Moreover, it is important to note that all implied volatility measures are sensitive to noise in option prices caused by market microstructure factors.

The early research on forecast performance comparison between Black-Scholes implied volatility and the time series methods came up with mixed results. For example, the results of Canina and Figlewski (1993) indicate that a naïve historical volatility measure forecasts the future volatility better than Black-Scholes implied volatility. Furthermore, Day and Lewis (1992) and Lamoureux and Lastrapes (1993) did not obtain unambiguous results, and were not able to make strong statement on the relative forecasting power of Black-Scholes implied volatility against forecasts generated with GARCH models. Even though the results of the early research on the subject are mixed, more recent studies, which attempt to correct various methodological problems of the earlier studies, found that the forecasting performance of Black-Scholes implied volatility is superior to the time series forecasts (see, for example Christensen & Prabhala 1998; Bluhm & Yu 2000; Ederington & Guan 2002; Jiang & Tian 2005). Interestingly, Christensen and Prabhala (1998) found that the forecast performance of Black-Scholes implied volatility improved after the stock market crash of 1987.

One of the most recent innovations in volatility forecasting literature is an implied volatility measure that does not rely on any specific model. A theoretical way to obtain such model-free implied volatility measure was presented by Britten-Jones and Neuberger (2000). This model-free implied volatility is not dependent on the restrictive assumptions of, for example, Black-Scholes model. Instead of relying on specific option pricing model, the volatility measure is based on the concept of arbitrage pricing, which lies in the center of modern option pricing theory. More specifically, the foundations of model-free implied volatility are in the research on the option implied return distributions. The research on the implied distributions was initiated by Breeden and Litzenbeger (1978) who presented a theoretically coherent way to extract underlying asset's risk-neutral return distribution from the option prices. Besides being a "model-free", their derivation is not subject to restrictive assumptions on e.g. asset return distribution. Rubinstein (1994), Derman and Kani (1994), Dupire (1994), and Jackwerth and Rubinstein (1996) made a further development on the subject, which ultimately led on the development of model-free implied volatility (Britten-Jones & Neuberger 2000, 839).

Britten-Jones and Neuberger (2000) reversed the procedure of the earlier research by taking a complete set of options as given, and used the prices of the options to extract information about the return generating price process for the underlying asset. They found that all processes consistent with the option prices generate a common risk-neutral expectation of the future volatility over a specified time period. In addition, they showed how to derive this model-free implied volatility. However, the implementation of this new volatility measure is not as straightforward as the option market conditions do not fully meet the requirements of the calculation procedure. Jiang and Tian (2005) were first to discuss the implementation of model-free implied volatility in practice, and they were first to test its forecasting performance as well.

Model-free implied volatility is theoretically superior to Black-Scholes implied volatility for two reasons. First, it is not subject to similar amount of restrictive assumptions as Black-Scholes implied volatility. Second, model-free implied volatility makes use of larger set of information than the typically used Black-Scholes implied volatility measures. For these reasons, model-free implied volatility should be a more accurate estimate for the market’s forecast of the future volatility. By being a more accurate estimate for the market’s forecast of the future volatility, model-free implied volatility should also be a better forecast for the ex post realized volatility as well. However, model-free implied volatility is not a perfect estimate for the market’s forecast. That is because it is estimated under the risk-neutral measure, which means that it does not incorporate variance risk premium. Hence model-free implied volatility should be a upwardly biased forecast for the ex post realized volatility.

Despite being a theoretically superior construct, model-free implied volatility’s implementation into practice is quite problematic. Due to the fact that option market conditions do not fully meet the requirements of the calculation of model-free implied volatility, a few compromises has to be made. Therefore, it is not clear if the forecast performance of model-free implied volatility is better than the forecast performance of Black-Scholes implied volatility, when the compromises regarding the implementation have been taken into account. The research on the subject is relatively limited and the results are mixed. While the results obtained by Jiang and Tian (2005) suggest that the model-free implied volatility indeed forecasts better than Black-Scholes implied volatility, Andersen and Bondarenko (2007) and Cheng and Fung (2012) came up with contrary results.

In conclusion, the vast majority of academic research on the volatility forecasting indicates that the implied volatility measures are better volatility forecasts than the forecasts created with time series methods. However, the research on the forecast performance of model-free implied volatility is still limited. Therefore, it is unclear how well model-free implied volatility performs compared to the time series methods and Black-Scholes implied volatility. This study aims to provide more information on this particular issue.

The results of this study suggest that model-free implied volatility contains useful information about the future level of latent volatility factor. However, despite the theoretical superiority of model-free implied volatility against Black-Scholes implied volatility, the results of this study also suggest that Black-Scholes implied volatility is superior volatility forecast compared to model-free implied volatility. This result applies for every measured aspect of forecast performance. Moreover, both implied volatility measures are found to outperform the forecasts generated with GARCH(1,1) model

1.2 Objective of the study

The objective of this study is to examine the forecasting performance of model-free implied volatility. The relative forecast performance of model-free implied volatility against some of the most common alternative methods, namely Black-Scholes implied volatility and GARCH(1,1), is under particular interest. The evaluation of forecast performance is done by constructing forecast series for the volatility of DAX index by using these three forecasting methods. DAX index is a German stock market index consisting of 30 largest companies traded in Frankfurt Stock Exchange. The options on DAX index are actively traded in EUREX derivatives exchange which facilitates the estimation of the implied volatility measures. The forecast series are constructed on weekly basis from September 2006 to September 2012 by using the set of information available one week prior to the time period of the forecast.

Four aspects of the forecast performance of model-free implied volatility are evaluated in this study. First of all, the information content of model-free implied volatility is analyzed by testing if it contains any information about the future volatility at all. Secondly, since theory suggests that model-free implied volatility is upwardly biased estimate for the future volatility, its biasedness and the biasedness of the competing models is tested. Thirdly, the forecast accuracy of model-free implied volatility is compared to the competing forecasts. Ultimately, the information content of the three forecasts is analyzed by testing which one of them contains most information about the future volatility. The forecasting performance of the model-free implied volatility is examined by testing four hypotheses that capture the above presented aspects of the forecasting performance. The hypotheses are following,

- Hypothesis 1: Model-free implied volatility contains information about the future volatility.
- Hypothesis 2: Model-free implied volatility is upwardly biased forecast for the future volatility.

- Hypothesis 3: Model-free implied volatility is more accurate forecast for the future volatility than Black-Scholes implied volatility or forecasts produced with GARCH(1,1).
- Hypothesis 4: Model-free implied volatility subsumes the information content (on the ex post realized volatility) of Black-Scholes implied volatility and forecasts produced with GARCH(1,1).

1.3 Structure of the study

The study begins with a presentation of the essential theoretical work on the subject. The theoretical part of the study is divided into three chapters. Chapter 2 presents the option pricing theory behind model-free implied volatility. In addition to an introduction to arbitrage pricing and risk-neutral valuation, chapter 2 presents Breeden and Litzenberger's (1978) findings on risk-neutral return distributions. Chapter 3 focuses on model-free implied volatility by presenting an intuitive theoretical background for it. Moreover, chapter 3 contains theoretical discussion on the suitability of model-free implied volatility for volatility forecasting purposes. Chapter 4 ends the theoretical part of the study by presenting the Black-Scholes option pricing model. Most importantly, theoretical work on the suitability of Black-Scholes implied volatility for volatility forecasting purposes is reviewed in chapter 4.

The issues related to empirical methodology are discussed in chapter 5. This is an essential part of the study, since the choice of the methodology is likely to have a significant effect on the results. In the first section of chapter 5, the implementation of model-free implied volatility is reviewed. The second section consists of discussion on the estimation of realized volatility series. Lastly, in the third section of the chapter the methods of forecast evaluation are under scrutiny.

The results of the study are presented in chapter 6, and chapter 7 consists of concluding remarks. Ultimately, chapter 8 summarizes the findings of the study.

2 OPTION PRICING THEORY BEHIND THE MODEL-FREE IMPLIED VOLATILITY

This chapter presents the theoretical framework behind model-free implied volatility. In order to understand the characteristics of model-free implied volatility, it is useful to understand the central concepts of the modern option pricing theory and the research on the option implied distributions, which are presented in this chapter. In the first section of the chapter, the concept of arbitrage is defined and its importance in modern option pricing theory is illustrated. The second and third sections of the chapter provide a brief introduction to risk-neutral valuation and Arrow securities, which are closely related methods for derivatives valuation, and are based on the concept of no-arbitrage. The last section of the chapter provides a view on the option implied distributions by presenting the risk-neutral density function derived by Breeden and Litzenberger (1978), which is particularly important, since the model-free implied volatility is derived from it.

2.1 Arbitrage-free markets

The concept of arbitrage is central to the modeling of financial markets which in turn enables us to price options. In modern finance, arbitrage is usually defined in more technical and narrower sense than traditionally. As defined in this financial sense, an arbitrage opportunity is a possibility to earn guaranteed profit without exposure to negative cash flows. In other words, it is possible to set up a portfolio, which may generate positive cash flows, but does not require investments, i.e. negative cash flows. In a financial marketplace, where the arbitrage opportunities exist and are exploited, it is impossible for the market to be in equilibrium. Modeling such chaotic market is practically impossible. In comparison, the absence of arbitrage is solely a sufficient requirement to build a relatively realistic model of the financial markets. Moreover, a model based on the no-arbitrage argument does not require assumptions on economic agents' behavior besides nonsatiation, i.e. they prefer more to less, or even more specifically, they always accept a cost-free increase in their consumption. The no-arbitrage argument is arguably the most important principle behind the modern option pricing theory. This is because a model of financial markets, building on no-arbitrage argument, makes the price determination for state contingent claims, e.g. options, possible. (Bingham & Kiesel 2004, 8-9)

Option pricing theory was heavily influenced by the breakthrough made by Black and Scholes (1973). They developed an option pricing model which is significantly different from the earlier models since it is only dependent on the observable variables. Black and Scholes assumed that the price of the underlying asset follows geometric

Brownian motion, and that the price of this asset and any option written on it are perfectly correlated. Given the possibility to borrow and lend at the risk-free rate, this setting enables us to create a portfolio that is a perfect substitute for the option. As it is possible to set up a hedging portfolio and the market is arbitrage-free, knowing the value of the underlying asset enables the valuation of the option. The difference between the earlier models and the Black-Scholes model is their remarkable insight: the expected return of the hedged position, consisting of the option and the corresponding hedging portfolio, has to be equal to the risk-free rate (Black & Scholes 1973, 640). This means that no assumption on the investors' risk preferences has to be made, since they are already embedded on the price of the underlying stock. In comparison, the earlier option pricing models required more assumptions on investors' behavior. For example, a model presented by Samuelson and Merton (1969) is dependent on the utility function that they assume for the marginal investor. More extensive presentation of Black-Scholes model is provided in Chapter 4.

The observations made by Cox and Ross (1976) generalized the findings of Black and Scholes (1973) and had a substantial impact on the pricing of all state contingent claims. In the Black-Scholes model, the equivalence of the option and the hedging portfolio was determined by the price process of the underlying security. Cox and Ross built a model consisting of one stock and one bond in an arbitrage free market, and showed that in this framework, value of any state contingent claim can be found by setting the stock return to risk-free rate and then computing the expected value of the claim. These results are robust to a relatively general class of price processes. The findings of Cox and Ross were further generalized by Harrison and Kreps (1979) and Harrison and Pliska (1981). These studies created a coherent theoretical foundation for risk-neutral valuation; a pricing method which is applicable for all state contingent claims.

2.2 Risk-neutral valuation

In general, if it is possible to replicate the payoff of an option using a portfolio of underlying securities in an arbitrage-free market, the price of the replicating portfolio must equal the price of the option. If the prices are not identical, an arbitrage opportunity exists, and the market is not arbitrage-free, i.e. it would be possible to earn a certain return with zero investment. As the value of the option and the hedging portfolio are equal in every possible state of the world, the no-arbitrage price is independent of the risk preferences of the investor. The construction of a portfolio that replicates the payoff of an option is a central problem in the option pricing theory. (Alexander 2008, 142)

In the event that it is possible construct a replicating portfolio for any possible option (or other contingent claim), the market is called complete. Bingham and Kiesel (2004) define the complete market as follows:

> *Given a set of financial assets on market, the underlying question is whether we are able to … replicate the cash flow of the new asset by means of a portfolio of our original assets. If this is the case and we can replicate every new asset, the market is called complete. (Bingham & Kiesel 2004, 22)*

Instead of pricing the option by identifying the replicating portfolio, it is possible to use risk-neutral valuation. The basic idea behind the risk-neutral valuation is that the option can be priced by discounting its expected cash flows at risk-free rate, since the investors' risk preferences are already embedded in the price of the underlying asset. This is because the risk that investor faces when investing in option, can be completely hedged in the case of complete markets. It means that any objective or subjective views that the investors have on the expected returns are irrelevant in option pricing, and the only assumption on investors' preferences is nonsatiation. (Bingham & Kiesel 2004, 115)

The risk-neutral valuation terminology has its roots in measure theory. The two widely used probability measures that are applied in asset pricing are the objective measure and the risk-neutral measure. Objective or the real-world measure $\boldsymbol{P}$ is determined by the investor's view about the future price of a risky asset. The future value of the asset is discounted to present value at risk-adjusted rate, where the investor's risk preferences determine the risk premium required. On the other hand, risk-neutral measure $\boldsymbol{Q}$ does not incorporate investors' risk preferences, which means that the future value is discounted to present at risk-free rate. Identifying the risk-neutral measure is an integral part of the risk-neutral valuation process. (Alexander 2008, 142)

Two conditions must hold for a probability measure to be a risk-neutral measure $\boldsymbol{Q}$. First, $\boldsymbol{Q}$ has to be equivalent to objective measure $\boldsymbol{P}$. The measures $\boldsymbol{P}$ and $\boldsymbol{Q}$ are equivalent, $\boldsymbol{P} \sim \boldsymbol{Q}$, if and only if they have the same null sets, i.e. the set of events having a positive probability is identical. Second, the underlying price process has to be a martingale under the risk neutral measure $\boldsymbol{Q}$. Broadly defined, a price process is martingale if the expected change of the process is zero. (Bingham & Kiesel 2004, 17, 43, 233)

An appropriate choice of numéraire is essential for the price process to be martingale under the risk-neutral measure. Numéraire is a French term meaning cash or currency, and in the language of finance, the numéraire is used to describe the unit of measurement for an asset that has value in specific currency. A probability measure of a random variable is defined with respect to the numéraire and a change of numéraire has no effect on the arbitrage opportunities. Since all investments return the risk-free rate under the risk-neutral measure, the expected growth in discounted asset prices is zero. Thus, setting a risk-free asset as a numéraire makes the price process martingale under the risk-neutral measure. (Alexander 2008, 144)

Furthermore, using the expectation operator to price options is based on the uniqueness of the risk-neutral measure, since it can be shown that its uniqueness is equivalent to the uniqueness of the no-arbitrage price for the state contingent claim. The existence and the uniqueness of the risk-neutral measure are dependent on two theorems.

- If the market is arbitrage-free, there exists at least one risk-neutral measure.
- If the market is complete, the risk-neutral measure is unique, which means that there exists only one risk-neutral probability for each state of the world.

Combining these two theorems, we get the first fundamental theorem of asset pricing, which states that in the arbitrage free complete market, there exists a unique equivalent martingale measure, which is the risk-neutral measure. (Bingham & Kiesel 2004, 118-119)

The Arrow securities are closely related to risk-neutral valuation, and they are important in option pricing, since all contingent claims can be priced in terms of Arrow securities. An Arrow security is a security that pays one unit of numéraire if a particular state of the world occurs, and nothing otherwise. The importance of these theoretical securities can be illustrated with the following example: if the payoff for a contingent claim that pays $F(s)$ if the state of the world is s, it can be replicated by buying $F(s)$ units of Arrow securities that pay one unit of numéraire if that particular state occurs. In an arbitrage-free market, the price of this replicating portfolio is equal to the price of the contingent claim. Due to the state contingent nature of these securities, the prices of Arrow securities are referred to as state prices. In order to price contingent claims in terms of Arrow securities, it is necessary to determine the state prices. (Arrow 1964; Sundaram 1997, 8)

In more general sense, options can be priced by multiplying its payoffs at each state by the corresponding state prices. The state prices and the risk-neutral probabilities are closely related: it can be shown that a state price discounted to present time at risk-free rate is equal to risk-neutral probability of that state. This result implies that option pricing with Arrow securities is equivalent to risk-neutral valuation, i.e. discounting the expected cash flows at risk-free rate. Furthermore, the first fundamental theorem of asset pricing can be written in terms of Arrow securities: the absence of arbitrage in a complete market is equivalent to the existence of a state price vector with unique positive value for every possible state. (Sundaram 1997, 9)

2.3 Risk-neutral return distribution implied by the option prices

Similarly to the Black-Scholes model, which can be used backwards to obtain the volatility implied by the options prices, it is possible to identify the risk-neutral density from the option prices. The risk-neutral density is defined as the market's objective estimate of the probability distribution for the level of the underlying asset's price on option expiration date. However, the risk-neutral density is not identical to the objective probability distribution, since it is modified by the variance risk premium when the objective probabilities are incorporated in option prices (Birru & Figlewski 2012, 152).

Breeden and Litzenberger (1978) presented a theoretically coherent way to extract the risk-neutral density from the option prices; given there is a complete set of options. The complete set of options means that there exists such set of options that there is an exercise price for every possible state of the world. Unlike the Black-Scholes implied volatility, the results of Breeden and Litzenberger are independent of restrictive distributional and variance assumptions. Instead, their results are solely based on the no-arbitrage argument.

The derivation of the risk-neutral measure is based on the no-arbitrage price of an elementary claim in terms of European call options. An elementary claim on any asset is defined as a security that pays $1 in T periods, if the value of certain asset is S at that time. If the value of the asset is not S in at that time, the elementary claim pays nothing. Breeden and Litzenberger (1978) showed that the payoff of an elementary claim can be replicated with short and long positions in call options. In an arbitrage-free market, the price of the replicating portfolio consisting of call options has to be equal to the price of the elementary claim. An elementary claim, as defined above, is practically identical to an Arrow security, and hence, given the connection between Arrow securities and risk-neutral probabilities, the risk-neutral density can be identified from the option prices.

Table 2.1 Call Option Payoffs

TABLE 2.1
Call Option Payoffs

S	$c(0,T)$	$c(1,T)$	$c(2,T)$
$S(T)=1$	1	0	0
$S(T)=2$	2	1	0
$S(T)=3$	3	2	1
⋮	⋮	⋮	⋮
$S(T)=N$	N	$N-1$	$N-2$

Breeden and Litzenberger (1978) set up a discrete model where the value of the portfolio in T periods has possible values of $M = \$1, \$2, \cdots, \$N$. The payoff vector of a European option with a strike price of X and expiration in T periods is denoted as $c(X,T)$. Two changes occur in the call option's payoff vector as its exercise price increases from X to $X+1$. First, with the set of states with $S=X+1$, the payoff becomes zero. Second, the payoffs in states where $S \geq X + 2$ are reduced by the change in exercise price. The payoffs in T periods for options with strike price of \$0, \$1 and \$2 are shown in the Table 2.1. (Breeden & Litzenberger 1978, 625)

In this setting, a portfolio with long and short position in call options with exercise prices of X and $X+1$ respectively, returns \$1 in every state where $S \geq X+1$. Similarly, portfolio with long and short position in call option with exercise prices of $X+1$ and $X+2$ respectively, returns \$1 in every state where $S \geq X+2$. The subtraction of these two portfolios, $[c(X,T)-c(X+1,T)]-[c(X+1,T)-c(X+2,T)]$, would have a payoff vector of, (Breeden & Litzenberger 1978, 625-626)

$$\begin{pmatrix}1\\1\\1\\\vdots\\1\end{pmatrix}-\begin{pmatrix}0\\1\\1\\\vdots\\1\end{pmatrix}=\begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix}.$$

The payoff vector indicates that the elementary claim for any given level of the market can be constructed with a portfolio consisting of call options. No-arbitrage argument implies that the price of the claim must equal the price of this portfolio. The portfolio paying \$1 only when the value of the asset is at S in T consists of one long call with $X=S-1$, two short calls with $X=S$ and one long call with $X=S+1$. (Breeden & Litzenberger 1978, 626)

In a more general setting where the step size between potential values of portfolio is ΔS, the price of a portfolio that pays \$1 if the value of the portfolio is S in T, periods is:

$$P(S,T;\ \Delta S)=(1/\Delta S)*\{[c(S-\Delta S,T)-c(S,T)]-[c(S,T)-c(S+\Delta S,T)]\}$$

Note that we can think of the set of $P(S,T;\Delta S)$ as state price-vector. If we divide the price of the portfolio $P(S,T;\Delta S)$ by ΔS, we get the price of the elementary claim in terms of call options:

$$\frac{P(S,T;\ \Delta S)}{\Delta S}=\frac{[c(S-\Delta S,T)-c(S,T)]-[c(S,T)-c(S+\Delta S,T)]}{(\Delta S)^2} \qquad (2.1)$$

And as the step size ΔS tends to zero:

$$\lim_{\Delta S\to 0}\frac{P(S,T;\ \Delta S)}{\Delta S}=\frac{\partial^2 c(X,T)}{\partial X^2}\bigg|X=S \qquad (2.2)$$

As the prices of the elementary claims can be seen as state prices Equations 2.1 and 2.2 can be interpreted as probability density functions for the underlying asset's future value; 2.1 in discrete case and 2.2 in continuous case. More precisely, it is the option market implied risk-neutral measure for marginal investor, since options are priced with risk-neutral valuation. (Breeden & Litzenberger 1978, 626)

An important feature of the option implied risk-neutral density is that it is independent of the price process for the underlying asset. The only requirements are that $c(X,T)$ is twice differentiable, there exists a complete set of options and the call option price $c(X,T)$ is strictly convex in the exercise price. If the $c(X,T)$ is not strictly convex in the exercise price, state prices are not strictly positive. As the options are priced with risk-neutral valuation, the implied distribution is not sensitive to investors' risk preferences and behavior as they are reflected in the call option prices. (Breeden & Litzenberger 1978, 627-628)

3 MODEL-FREE IMPLIED VOLATILITY

This chapter consists of an introduction to model-free implied volatility and theoretical discussion on its suitability for volatility forecasting purposes. The chapter is divided into two sections. First section follows the presentation of Britten-Jones and Neuberger by giving an intuitive background for the derivation of model-free implied volatility, while the latter section of the chapter consists of discussion on how the risk-neutrality of model-free implied volatility affects its use in volatility forecasting.

3.1 Introduction to model-free implied volatility

Traditionally, the use of specific option pricing model has been prerequisite for the extraction of implied volatility measure from the option markets. However, this approach is subject to several shortcomings. Most importantly, when the model-based approach is implemented, the comparison between the realized and implied volatility is a joint test of the option pricing model and the information efficiency of the option market. Britten-Jones and Neuberger (2000) made a significant departure from the earlier research by presenting a way to derive implied volatility from the option market in a way that does not require a specific model. Instead, their implied volatility measure is based on the findings of Breeden and Litzenberger (1978), which in turn builds on the no-arbitrage argument. Using the results of Breeden and Litzenberger as a premise, Britten-Jones and Neuberger discovered that the risk-neutral integrated return variance between current date and a future date is fully specified by a complete set of option prices.

Jiang and Tian (2005) presented an alternative, simpler, way to derive the model-free implied volatility, and generalized the results of Britten-Jones and Neuberger to all martingale processes. Additionally, they showed a way to implement the model-free implied volatility in practice, and tested its information content. They found that the forecasting power of model-free implied volatility is informationally superior to Black and Scholes implied volatility and the forecasts produced with time series models. However, Cheng and Fung (2012) and Andersen and Bondarenko (2007) came up with a contrary result when they tested the forecasting power of the model-free implied volatility against various alternative volatility forecasting methods. They found that Black-Scholes implied volatility forecast subsumes the information contained in model-free implied volatility.

The model-free implied volatility is obtained from the market prices as the amount that the market is willing to pay in order to receive the sum of squared returns between current time and an arbitrary date in future (Britten-Jones & Neuberger 2000, 847). This can be seen as the market's prediction of the future volatility since the squared returns are typically used as a measure of volatility. Additionally, it is important to note that this value is a forward price, which has interesting consequences. Since model-free implied volatility is calculated under the risk-neutral measure, it is discounted to present value at risk-free rate, and, hence, in the presence of variance risk, it differs from the corresponding objective value. Due to the difference between risk-neutral and objective values, the model-free implied volatility is not pure volatility forecast as it excludes the premium associated with variance risk. Even so, the model-free implied volatility is likely to be closely correlated with the future realized volatility (Andersen & Bondarenko 2007, 1).

In order to understand the intuition behind the model-free implied volatility, it is necessary to understand some characteristics of conditional volatility. Britten-Jones and Neuberger (2000) presented a discrete setting where, building on the pioneering work of Breeden and Litzenberger (1978), the risk-neutral measure for the stock price movements can be extracted from option prices. Additionally, they study the price processes implied by the risk-neutral measure, and conclude that knowing the risk-neutral probabilities for the price movements does not give sufficient amount of information to identify the price process for the underlying asset.

Britten-Jones and Neuberger (2000, 841-842) used a setting where both time and stock price are discrete. However, the results are not driven by the discrete setting as they can be generalized for continuous setting as well. The discrete time has an interval of h and is a finite set of times $\boldsymbol{T}$, ranging from 0 to T.

$$\boldsymbol{T} = \{0, h, 2h, \cdots, T\}$$

The initial price of the stock is denoted by S_0 and it can take values in set $\boldsymbol{K}$, where $\boldsymbol{K}$ is a finite geometric series of possible stock prices. Furthermore, the stock price is assumed to be continuous in the sense that each in period, it can only move up or down one level, or stay at the same level. The continuity assumption does not impose arbitrary restriction on the volatility, since, by making the time interval shorter, volatility can be raised without a limit. (Britten-Jones & Neuberger 2000, 842-843)

$$\boldsymbol{K} = \left\{K : K = S_0 u^i, i = \pm 1, \pm 2, \cdots, \pm M\right\}, where\ u > 1$$

$$if\ |i - j| > 1, then\ Pr\left\{S_{t+h} = S_0 u^j \middle| S_t = S_0 u^i\right\} = 0$$

At the time zero, there exists a complete set of European call options with prices $c(K, t)$ for all $t \in \boldsymbol{T}$ and $K \in \boldsymbol{K}$. Dividends and the interest rates are assumed to be zero. This assumption does not lead to loss in generality, since in the presence of nonzero interest rates and dividends the prices of the option and the underlying asset can be viewed as forward prices, which can be converted back to spot prices. (Britten-Jones & Neuberger 2000, 842)

As stated above, the absence of arbitrage in a complete market is equivalent to the existence of a unique risk-neutral measure. This in turn enables us to use risk-neutral valuation which means that the price of the security is the expected value of its discounted payoffs. This setting allows us to determine the risk-neutral probability for the stock price to reach certain level. However, the price path leading to a certain price is not determined, since the probability is only conditioned on the last price instead of whole price history. (Britten-Jones & Neuberger 2000, 842)

The risk-neutral probability of the stock price to reach any two price levels at two consecutive periods is determined in terms of the initial option prices. The intuition is easily demonstrated by considering a portfolio with short position in European call option expiring at t and long position in European call option expiring at $t+h$. Both of these options have a strike price of K. The value of this position is $c(K,t+h) - c(K,t)$. Britten-Jones and Neuberger (2004, 858) derived a hedging strategy which results in a zero-payoff unless the stock price is K at t and Ku at $t+h$ in which case the payoff is $K(u-1)$. As the risk-neutral valuation states, the expected payoff of a portfolio equals its cost. This gives us the risk-neutral probability for an upward movement of the stock price. From this result, it is possible to derive the risk-neutral probabilities for stock price to move down or stay at the same level: (Britten-Jones & Neuberger 2000, 842):

$$Pr\{S_t = K \; and \; S_{t+h} = K^*\}$$

$$= \begin{cases} \dfrac{C(t+h,K) - C(t,K)}{K(u-1)}, if \; K^* = Ku \\ \dfrac{C(t,Ku) - (1+u)C(t+h,K) + uC(t,K/u)}{K(u-1)}, if \; K^* = K \\ \dfrac{C(t+h,K) - C(t,K)}{K(1-1/u)}, if \; K^* = K/u \\ 0, otherwise \end{cases} \quad (3.1)$$

The joint probability can be written as a product of marginal and conditional probabilies:

$$Pr\{S_t = K \; and \; S_{t+h} = Ku\} = Pr\{S_t = K\} * Pr\{S_{t+h} = Ku | S_t = K\} \quad (3.2)$$

If we substitute the result from equation 3.1 to equation 3.2 and rearrange we can solve the conditional probability of equation 3.2 in terms of the option prices:

$$Pr\{S_{t+h} = Ku | S_t = K\} = \frac{C(t+h,K) - C(t,K)}{C(t,Ku) - (1+u)C(t,K) + uC(t,K/u)}$$

The conditional probabilities for the stock price to move down or stay at the same level are given by the martingale condition and the fact that the probabilities must sum to one. Now, the conditional expectation of the squared return can be calculated. The following result applies for any continuous risk-neutral process: (Britten-Jones & Neuberger 2000, 844)

$$E\left[\left(\frac{S_{t+h} - S_t}{S_t}\right)^2 \middle| S_t = K\right] = \frac{[C(t+h,K) - C(t,K)](u-1)^2\,(u+1)/u}{C(t,Ku) - (1+u)C(t,K) + uC(t,K/u)} \quad (3.3)$$

Equation 3.3 can be interpreted as volatility forecast for the stock price conditional on the last price. The unconditional expected squared return between two arbitrary dates t_1 and t_2 can be derived from this result: (Britten-Jones & Neuberger 2000, 846)

$$E\left[\sum_{t\in[t_1,t_2-h]} \left(\frac{S_{t+h} - S_t}{S_t}\right)^2\right] = (u - 1/u) \sum_{K\in\boldsymbol{K}} \frac{C(t_2,K) - C(t_1,K)}{K}$$

As the time interval h and the step size $u-1$ tends to zero, and the t_1 is set to current time 0, we get the following formula: (Britten-Jones & Neuberger 2000, 846)

$$E\left[\int_0^{t_2} \left(\frac{S_{t+h} - S_t}{S_t}\right)^2\right] = 2\int_0^{\infty} \frac{C(t_2,K) - max(S_0 - K, 0)}{K^2} dK$$

It is important to note that the above forecast is made under risk-neutral probabilities, which may differ from the objective probabilities, and, the forecast is a forward price, since the interest rates are assumed to be zero. In the presence of positive interest rates, the model-free implied volatility is defined as: (Jiang & Tian 2005, 1314)

$$E\left[\int_0^{T} \left(\frac{dF_t}{F_t}\right)^2\right] = 2\int_0^{\infty} \frac{C^F(T,K) - max(F_0 - K, 0)}{K^2} dK, \quad (3.4)$$

$$F_t = S_t e^{r(T-t)},$$

$$C^F(T,K) = C(T,K)e^{r(T-t)}.$$

Equation 3.4 can be interpreted as an unconditional volatility forecast from current time to date T (Britten-Jones & Neuberger 2000, 846). Britten-Jones and Neuberger (2000) and Jiang and Tian (2005) present a proof for this result. Intuitively, the volatility forecast can be seen as the market price for the sum of squared returns, implied by the option prices. The use of sum of squared returns to proxy sample variance is a common practice in variance estimation. The continuity assumption implies that the volatility forecast is valid when the stock price follows diffusion process. Britten-Jones and Neuberger (2000) do not discuss the validity of the result in a case where the underlying price process contains jumps. However, Jiang and Tian (2005) generalized the model-free implied volatility, by showing first that the forecast measure is valid for price process that contains jumps and further showing that it is valid for all martingale processes. Thus the forecast measure is valid for an especially general class of price processes.

3.2 The bias in risk-neutral volatility forecast

Due to the fact that model-free implied volatility is calculated under the risk-neutral measure, it is not a pure volatility forecast for the underlying asset. In the presence of variance risk premium, the risk-neutral and the objective measures are not identical. The variance risk premium is the premium that rational risk-averse investor requires for the variation in variance. Hence, such investor is willing to pay higher price for an asset with constant volatility compared to otherwise identical asset with stochastic volatility. As expected, evidence from the option pricing literature indicates that the market price of stochastic variance is negative which means that the marginal investor requires a positive premium for carrying variance risk. (Jiang & Tian 2005, 1327)

As the underlying asset is priced under the objective measure, it is not clear if the model-free forecast is relevant in the real world. Unlike the objective probabilities, risk-neutral probabilities do not incorporate investors’ risk preferences. Since the model-free volatility forecast is a forward price, it is biased under the objective measure in the presence of volatility risk premium. Generally the volatility forecast made under the risk-neutral probabilities is upward biased estimate for the realized volatility as the marginal investor is not willing to pay as much as the forecast suggests in order to receive the sum of squared returns, since the market price of stochastic volatility is negative. This means that the size of the risk premium required by the marginal investor is an important factor when analyzing the relevancy of model-free implied volatility forecast.

The variance risk premium can come from two different sources. First possible source is that it comes from the correlation between variance risk and the asset return, while the second possible source is that it is a separate premium for the asset’s variance variation, which is independent of the return risk premium (Carr & Wu 2009, 1312).

Carr and Wu (2009) developed a direct and robust method for quantifying risk-premium for stochastic volatility. Their methodology is based on determining synthesized variance swap rate. The variance swap is a derivative security which pays the difference between a standardized estimate for the realized volatility and the variance swap rate. The variance swap rate is the rate that makes the initial value of the contract zero. Carr and Wu showed that the synthesized swap rate can be determined using the prices in the option market. Since the swap contract is priced under the risk-neutral measure, they proposed that the difference between ex post realized volatility and the synthetic variance swap rate is a valid estimate for the variance risk premium. Their analysis shows that there exists a significant negative risk premium for stochastic volatility. Additionally, they concluded that the correlation with the asset's return explains only a small portion of the variance risk premium, and the majority of the variance risk premium is generated by a risk factor that is independent from asset's variance variation. (Carr & Wu 2009, 1311-1312; 1338)

Bakshi and Madan (2006) attempted to find the theoretical determinants of the volatility spread, i.e. the difference between the objective and the risk-neutral volatilities. Their empirical results on the relationship between the return's higher order central moments and the volatility spread, applicable for a relatively general class of pricing kernels and market return dynamics, suggests that the volatility spread is more distinct when the objective measure is leptokurtic and left-skewed. In other words, the variance risk-premium increases when the investors' anticipate the risk of extreme losses more severe. Overall, Bakshi and Madan (2006, 1955) conclude that the volatility spread is systematically related to the higher moments of the objective distribution.

The research on the option implied measures suggests that there is a link between the two measures, which enables us to perform a transformation between the risk-neutral and the objective measure. However, the transformation requires several assumptions in order to be valid. For example, the representative utility function has to be identified (Liu, Shackleton, Taylor & Xu 2007, 1504). Rubinstein (1994, 803-804) compared the shape of the risk-neutral and objective distributions. Based on an assumption that the representative utility function is logarithmic, his analysis indicates that change of measure from risk-neutral to objective does not induce a significant change in the shape of the distribution. The change of measure moves the distribution to right by the amount of variance risk premium, but the qualitative shape of the distribution remains practically unchanged. The results are robust to another concave utility function tested by Rubinstein.

In conclusion, in order to the model-free volatility forecast to be unbiased under the objective measure, the variance risk is required to be unpriced. It is an established fact in the financial marketplace that the return variance is stochastic and has a negative price, which implies that the model-free forecast is an upward biased estimate for the realized volatility. The prior research on the subject offers two important results which should be considered when analyzing the forecasting power of the model-free implied volatility. Firstly, risk-neutral forecast is often considered relevant in the real world, since the upward bias of the forecast is generally considered negligible (Britten-Jones & Neuberger 2000, 847). Secondly, the bias is likely to be higher when the financial markets expect the risk of extreme losses to be more severe (Bakshi & Madan 2006).

4 BLACK-SCHOLES IMPLIED VOLATILITY

The Black-Scholes implied volatility is a by-product of the widely used theoretical option pricing model presented by Black and Scholes (1973). The implied volatility, extracted from the option prices with Black-Scholes model, is generally viewed as the market-based volatility forecast, and hence it is widely considered to be superior to time series volatility forecasts (Poon 2005, 115). This is because Black-Scholes implied volatility is derived from the option prices, and hence it is considered to contain the information that the market participants use in the price formation in the financial marketplace. This argumentation implicitly assumes that the Black-Scholes model is the "correct" model for option pricing, at least to the extent that it is able to capture some incremental information on the future volatility from the prices. Moreover, for the argumentation to be valid, the market participants are required to possess more information about the future volatility than the time series models can provide. In order to understand the characteristics of the Black-Scholes implied volatility, it is useful to understand the key principles and the assumptions behind the model.

4.1 Black-Scholes option pricing model

The Black-Scholes option pricing model has reached significant position in the option pricing theory. It has become the benchmark valuation model in academic literature, and, impressively, in the use of practitioners as well (Figlewski 2008, 1). Hence, the model is often regarded as one the most successful and most widely used in the social sciences (Rubinstein 1994, 772). One of the main reasons for the attractiveness of the Black-Scholes pricing equation is that it is a general equilibrium formulation of the option valuation problem. Moreover, it is straightforward to implement in practice, since most of its parameters are observable in the financial marketplace. The underlying principle behind the model is that, if the options are priced correctly in the market, it should not be possible to find arbitrage opportunities by creating short and long positions in options and their underlying assets. Using this reasoning, Black and Scholes (1973) derived a theoretical pricing equation for options.

Even though the Black-Scholes model is closely related to earlier models presented in the option pricing literature (cf. Sprenkle 1961; Samuelson & Merton 1969), there is one vital difference between them: Black and Scholes derived the theoretical price for option in such manner, that the price is not dependent on the risk preferences of investors, but only on observable variables. However, Black-Scholes model is based on a set of relatively restricting assumptions, as it assumes the market conditions to be "ideal".

The assumptions required for the Black-Scholes pricing equation to be valid are: (Black & Scholes 1973, 640)

- The short term interest rate (risk-free rate) is known and constant. However, this assumption is not necessary, since Merton (1973, 167) derived the model under weaker assumptions, and showed that the pricing equation is still valid if the interest rate is stochastic.
- The instantaneous return for the underlying asset is described by the geometric Brownian motion, implying that the distribution of possible prices for the underlying asset is lognormal and its return volatility is known and constant.
- The market conditions are frictionless: there are no transaction costs or differential taxes, trading takes place continuously and it is possible to lend and borrow at the risk-free rate.
- There are no dividend payments before the option's date of maturity. However, there are modifications of the model that allow dividend payments (cf. Merton 1973).

Black and Scholes (1973, 641) show that under these assumptions, the theoretical price of the option is only dependent on the on initial price of the underlying asset and variables that are known and constant. The model's assumptions allow us to create a hedged portfolio, consisting of long position in the underlying asset and short position in the option. The value of this portfolio is not dependent on the price of the underlying asset, but is only dependent on time and the known constant variables. In order to maintain the hedged position, the amount of assets and options to be hold changes as the time passes and the price of the underlying changes. If the hedge is maintained continuously, the return of the hedged position is completely independent of the price of the underlying asset. As the return on the hedged portfolio is certain, in an arbitrage-free market, it must be equal to the risk-free rate. Black and Scholes (1973) derived the price for the option from the following equation implied by the no-arbitrage argument: the hedged portfolio must return the risk-free rate. It is important to note that the expected return for the underlying asset is irrelevant and thus does not appear in the pricing equation. Under these assumptions, the theoretical pricing equation for a European call option takes following form (Black & Scholes 1973, 644):

$$C(S,t) = \phi(d_1)S - \phi(d_2)Ke^{-r(T-t)}, \qquad (4.1)$$

$$d_1 = \frac{ln\left(\frac{S}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}},$$

$$d_2 = \frac{ln\left(\frac{S}{K}\right) + \left(r - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}.$$

Where,
$\phi(\cdot)$ is the cumulative standard normal distribution function,
S is the spot price of the underlying asset,
K is the strike price of the option,
r is the risk-free rate,
σ is the expected return volatility for the underlying asset from current time to date of expiration, and,
$T - t$ is the time to date of expiration.

The price of the underlying asset S follows geometric Brownian motion with drift μ and volatility σ:

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

Where,
W_t is a Wiener process.

Although Black-Scholes model does not require information about the investors' risk preferences, its assumptions are quite restrictive. The model assumes the underlying asset's price process to follow geometric Brownian motion which determines its return distribution and sets restrictions on its return volatility. The Wiener process determines the return distribution of the underlying asset to be lognormal. The lognormal return distribution is inconsistent with the empirical evidence which indicates that typical return distribution for financial assets is leptokurtic and left-skewed (cf. Rubinstein 1994). Moreover, the volatility parameter in the geometric Brownian motion is constant, even though it is a well-documented fact that financial assets' return variance is stochastic, (cf. Carr & Wu 2009, 1311). Due to the restrictive assumptions, Black-Scholes model can be considered as an approach used in academia and in practice rather than optimal behavior in the financial market with stochastic volatility (Canina & Figlewski 1993, 661). Moreover, the theoretical Black-Scholes option price does not incorporate the supply and demand factors which, after all, determine the actual prices in the marketplace.

One of the particularly attractive features of the Black-Scholes model is that almost all of its parameters are directly observable from the financial marketplace. The only parameter that has to be forecast is the return volatility of the underlying asset over the option's maturity. As the volatility in the financial market tends to be stochastic, forecasting it is a difficult task. Given there is financial marketplace with price quotations for options written on the same underlying asset and with identical date of expiry, the forecasting problem can be solved by using the Black-Scholes pricing equation backwards: it is possible to find the level of future volatility over option's life time that makes the option's value consistent with the prices observed in marketplace. This level of future volatility, consistent with option's price obtained with Black-Scholes model, is often referred to as the Black-Scholes implied volatility. (Canina & Figlewski 1993, 659; Figlewski 2008, 1)

Two characterizing results for the Black-Scholes implied volatility can be derived from the pricing equation. Firstly, the implied volatility is always greater than zero and, secondly, there is a clear correspondence between implied volatility and option price as the price of the option increases as the implied volatility increases. (Poon 2005, 73-74)

$$\sigma_{BSIV} > 0$$

$$\frac{\partial C_{BS}}{\partial \sigma_{BSIV}} > 0$$

The implied volatility obtained with the backward working use of Black-Scholes is often interpreted as a market's forecast for the underlying asset over the option's maturity. The reasoning behind this is that the option prices reflect the market participants' expectations about the future movements of the underlying asset. Since the Black-Scholes implied volatility makes use of comprehensive set of information observable in the marketplace, it should be informationally superior to forecasts obtained with time series models. Hence, Black-Scholes implied volatility is commonly used in volatility forecast purposes, regardless of the model's unrealistic assumptions. (Poon 2005, 115)

4.2 Volatility smiles and smirks

In most option markets, numerous options, identical with respect to expiry and underlying asset, are traded at various different strike prices. When the Black-Scholes model is applied inversely, each of these options yields an implied volatility for the underlying asset over option's maturity. In Black-Scholes world, all of these implied volatilities should be identical, since they forecast volatility for the same asset over the same period. However, in real world, it is an established fact that the Black-Scholes implied volatilities differ across the strike prices. When the implied volatility is plotted against strike price, the shape of the graph typically forms a skewed pattern. Although some differences in the Black-Scholes implied volatility across the strike prices tend to be persistent, others, specifically those due to market microstructure factors such as sparse trading, bid-ask spreads and discrete prices, are typically transitory. (Ederington & Guan 2002, 812; Poon 2005, 74)

Empirical evidence shows that, before the stock market crash of 1987, the Black-Scholes implied volatility plotted against strike prices was often observed to be U-shaped, with minimum value located near at-the-money options. In option pricing literature this pattern is referred to as the volatility smile. After the stock market crash of 1987, the typically observed connection between the implied volatility and option moneyness changed to a less symmetric, skewed, pattern. Compared to the pre-crash U-shaped volatility pattern, the post-crash implied volatilities tend to be higher for options with especially low strikes, and lower for options with high strikes. This post-crash pattern is often referred to as volatility smirk or skew. (Poon 2005, 74)

In the Black-Scholes model, the price of the option increases as the implied volatility increases. Hence, the volatility smiles and smirks imply that there is a premium required for options with strike price far from the current price of the underlying asset. The standard explanation for this is that the data from the financial markets differs from the Black-Scholes' assumptions in two systematic ways: first, the return variance is not constant, and, second, the lognormal return distribution typically underestimates the chance for large movements in the underlying price. For example, if the market expects the probability of the extreme losses to be higher than the lognormal return distribution implies, the probability of exercise for deep-out-of-the-money call option is higher than Black-Scholes pricing equation indicates. Thus the market prices for deep-out-the-money call options are higher than they would be in Black-Scholes world, and the higher prices produce higher implied volatilities. (Poon 2005, 75-76)

The empirical evidence supports the explanation that the unrealistic distribution assumption is the reason behind the observed smile and smirk patterns. Jackwerth and Rubinstein (1996, 1628-1629) studied the implied distribution for S&P 500 index, and found that, before the crash of 1987, the option implied return distribution tended to be slightly left-skewed and platykurtic. That is, the mode is located left from the mean and the mode is less pronounced compared to lognormal distribution. After the crash of 1987, the option implied return distribution changed significantly as it become even more left-skewed and it changed from platykurtic to leptokurtic. In other words, the market participants' expect the probability of a plunge in index value significantly higher than they did before the crash. The observations of Jackwerth and Rubinstein are consistent with the change in typically observed pattern in Black-Scholes implied volatilities.

4.3 Black-Scholes implied volatility in forecasting

The forecasting performance and the information content of the Black-Scholes implied volatility is an extensively studied area of research. The early studies on the forecasting power of the Black-Scholes implied volatility found that it contains little incremental information beyond time series forecasts (see, for example Canina & Figlewski 1993). However, the early research the on the topic was subject to a number of methodological shortages, and the more recent research indicates that Black-Scholes implied volatility is significantly more efficient forecast for the realized volatility compared to the forecast's obtained with time series models. Hence, the empirical research indicates that the implied volatility contributes a statistically significant amount of information about the future realized volatility. (Jiang & Tian 2005, 1305-1306)

As in the model-free implied volatility, the Black-Scholes implied volatility is a biased estimate for the future realized volatility. The bias in the Black-Scholes implied volatility has a different source than the bias in model-free implied volatility, because in the Black-Scholes world there is no variation in variance as it is assumed constant. This implies that there is no variance risk premium and the expectations of the squared returns are identical under the objective and risk-neutral measures. Moreover, it can be shown that the risk-neutral expectation of the squared volatility is the square of Black-Scholes implied volatility. However, the unbiased estimate applies only for the squared form of the volatility, and the implied volatility itself is a biased estimate for the future realized volatility. This is due to Jensen's inequality which indicates that the implied volatility forecast for volatility itself is a positively biased under the risk-neutral measure. Even in the case where the implied volatilities form a flat pattern, the Jensen's inequality holds strictly. Unless the volatility is actually constant as assumed in the Black-Scholes model, the inequality holds. (Britten-Jones & Neuberger 2000, 846-847)

Since options with different strike prices tend to yield different implied volatilities, it is not evident, which implied volatility or combination of implied volatilities is the most accurate forecast for the future volatility over the life of the option. The prevailing practice is to use either at-the-money (ATM) implied volatility or an averaged implied volatility for the forecasting purposes. The use of ATM options in justified with the fact that they generally have the highest trading volume and they are most sensitive to the volatility input. When the ATM option is not available, which is commonly the case in the practical applications, the nearest-to-the-money (NTM) option is used instead. (Poon 2005, 116-117)

The average used in forecasting can simply be the equally weighted average of the implied volatilities or a product of a more complex weighting scheme. Vatious different weighting schemes have been introduced by both the academics and the commercial vendors. Ederington and Guan (2002) tested the forecasting performance of some of the most popular weighting schemes by studying the implied volatilities produced by S&P 500 futures options. They found that in the case of S&P 500 futures options, the choice of weighting scheme is of negligible importance. This is because the amount of noise in implied volatilities is close to non-existent, i.e. each of the implied volatilities yields virtually the same forecast if the bias is corrected. Additionally, they found that the bias in Black-Scholes implied volatility is relatively stable through time, which means that it is possible to adjust the forecast to correct at least some of the bias. Nevertheless, the importance of the weighting scheme could be greater with less liquid options, which prices contain more noise.

5 EMPIRICAL METHODOLOGY

This chapter consists of discussion on the issues related to the empirical methodology. The issues covered in this chapter are the implementation of model-free implied volatility, estimation of the realized volatility and the forecast evaluation criterion. A number of ways to solve the relevant empirical issues are proposed in the volatility forecasting literature, which is why it is necessary to find the methods that are most suitable for this study. The decisions on methodological issues are of great importance since they may have significant effect on the results of the study. Moreover, the data and the methodology implemented in this study are presented in the final section of this chapter.

5.1 Implementation of model-free implied volatility

In this section, the implementation of the model-free implied volatility is discussed. The discussion on the implementation procedure of model-free implied volatility follows closely the methodological issues presented by Jiang and Tian (2005). As stated in chapter 3, in the presence of positive interest rates, the model-free implied volatility is defined as:

$$E\left[\int_0^T \left(\frac{dF_t}{F_t}\right)^2\right] = 2\int_0^\infty \frac{C^F(T,K) - max(F_0 - K, 0)}{K^2} dK \quad (5.1)$$

The model-free implied volatility is defined as an integral of a continuous set of call option prices. In order to calculate the model-free implied volatility with numerical integration methods, a continuum of strike prices is required. However, in practice the options are traded on a finite number of discrete strikes and the market prices for options are likely to contain noise caused by market microstructure factors such as sparse trading and bid-ask bounces. Relatively small irregularities in option prices may cause serious irregularities in the risk-neutral probabilities (Figlewski 2008, 2). Even if there is an extensive amount of strikes available, trading typically takes place around at-the-money level, which means that the prices of the deep-in- or deep-out-of-the-money options are more likely to contain microstructure noise. The trading is especially thin on in-the-money options (Cheng & Fung 2012, 796).

5.1.1 Issues with the numerical integration

Jiang & Tian (2005, 1309) discuss the truncation errors, which are present when the model-free implied volatility is calculated with numerical integration methods and the trading takes place at a number of strike prices in certain interval. That is, the tails of the distribution for the underlying asset are not covered by the set of available strikes. Jian and Tian consider a setting, where interval $[K_{min}, K_{max}]$, $0 < K_{min} < F_0 < K_{max} < \infty$, specifies the range of available strike prices. In order to concentrate on the truncation errors, they assume a continuum of strikes inside the defined interval. In this setting, the model-free implied volatility is calculated as an integral between the minimum and maximum strike prices (Jiang & Tian 2005, 1309):

$$2\int_{K_{min}}^{K_{max}} \frac{C^F(T,K) - max(0, F_0 - K)}{K^2} dK$$

The truncation errors are present when the tails of the distribution for the underlying asset are not covered by the interval $[K_{min}, K_{max}]$. Jiang and Tian (2005, 1310-1311) examined the errors caused by truncation by generating a continuum of call prices with stochastic volatility model. As a result of this analysis, they found that the truncation error declines monotonically as the end point of the truncation interval moves away from the current price of the underlying asset, and that it becomes negligible if the end points of the truncation interval are more than two standard deviations away from the future price of the underlying asset. Since the DAX index options are traded on a relatively wide interval of strike prices, the truncation errors caused by truncation are likely to be insignificant in this study.

In addition to truncation errors, the use of numerical integration introduces error caused by discretization. Jiang and Tian (2005, 1310) propose the use of trapezoidal integration for numerical calculation of the model-free implied volatility:

$$2\int_{K_{min}}^{K_{max}} \frac{C^F(T,K) - max(0, F_0 - K)}{K^2} dK \approx \sum_{i=1}^{m} [g(T, K_i) + g(T, K_{i-1})]\Delta K$$

Where,
$\Delta K = (K_{max} - K_{min})/m$,
$K_i = K_{min} + i\Delta K$ for $0 \leq i \leq m$,
$g(T, K_i) = [C^F(T, K_i) - max(0, F_0 - K_i)]/K_i^2$. (Jiang & Tian 2005, 1313)

If the amount of subintervals is assumed finite, $m < \infty$, the results of numerical integration are subject to discretization errors, when the trapezoidal rule is used. The error in trapezoidal integration is defined as:

$$E \leq \frac{\Delta K^3}{12} mF$$

Where,

F is the maximum absolute value of the second derivative in the interval ΔK. (Phillips 2003, 126)

As in the case of truncation errors, Jiang and Tian (2005, 1313) studied the discretization errors in model-free implied volatility with stochastic volatility and random jump model. The interval ΔK was measured in terms of standard deviation in order to generalize the results. They concluded that the discretization errors are negligible if the interval ΔK is shorter than $0{,}35$ times the standard deviation. In conclusion, according to Jiang and Tian (2005) the errors caused by numerical integration are likely to be insignificant.

5.1.2 Issues with the limited availability of strike prices

Inside the sample range, the call options on DAX index are traded at an interval of 50 index points. Without the assumption that there is a continuum of strikes between the defined range $[K_{min}, K_{max}]$, the model-free implied volatility has to be extracted from a sparse set of discrete strikes. According to Jiang and Tian (2005, 1315), a curve-fitting method is the most effective way to deal with this issue, since it is easy to implement in practice. However, the use of curve-fitting methods to option prices is problematic as there is a nonlinear relationship between the option price and strike price. The nonlinear relationship causes numerical difficulties when the curve-fitting methods are applied.

A typical way to solve this problem is to transfer the option prices into implied volatilities and apply the curve-fitting methods to implied volatilities (cf. Shimko 1993). The prices of the call options are first converted into Black-Scholes implied volatilities, and then the curve fitting methods are applied to implied volatilities. After the interpolation, the implied volatility curve is translated back in to call option prices using the Black-Scholes pricing equation. Since the Black-Scholes model is only used as a computational device, this procedure does not require the Black-Scholes assumptions to be valid. If the Black-Scholes assumptions would be valid, all of the options would yield identical implied volatilities. If the interpolated implied volatilities would be transferred back to prices in Black-Scholes world, the resulting risk-neutral density would be well-behaved and lognormal. Despite the unrealistic assumptions of the pricing model, the smile and smirk patterns produced by Black-Scholes implied volatilities tend to be relatively smooth and well-behaved, which facilitates the interpolation. (Figlewski 2008, 11-12)

Jiang and Tian (2005, 1316) make use of cubic spline interpolation method to fit a smooth function of implied volatilities. However, the use of cubic spline – typically the first choice for interpolation – is problematic in the case of the implied volatilities. Cubic spline generates a curve that is forced to go through every implied volatility observation, which means that the price irregularities caused by the market microstructure factors affect the shape of the curve (Figlewski 2009, 12). Risk-neutral density extracted from such curve may contain negative probabilities and other serious irregularities. The advantage of the cubic splines method is that it is smooth across range and provides an accurate fit to known implied volatilities (Jiang & Tian 2005,1316).

The use of cubic spline method is only effective for fitting curve for the implied volatilities in the sample range. For implied volatilities with strike prices that are higher or lower than the available range, the endpoint observation has to be used to extrapolate the option values. This implies that the volatilities are assumed to be constant outside the sample range and hence the tails of the distribution are lognormal (Figlewski 2008, 17). Fitting the lognormal tails for the risk-neutral density introduces approximation error that is different from the truncation errors. Jiang and Tian (2005, 1316) compared the errors caused by truncation and extrapolation, and found that it is generally more effective to fit the tails by using extrapolation. The truncation method ignores the strike prices outside the sample range, and hence underestimates true volatility, whereas extrapolation solves the underestimation problem by assuming implied volatility function to be flat outside the sample range.

Although the approximation error is less severe when the extrapolation is used, it is still present with the typically observed implied volatility patterns. For example, if the observed implied volatility is smile shaped, the implied volatility is increases as the strike price increases or decreases from the current price for the underlying asset. Hence, assuming the implied volatility function constant outside sample interval underestimates the out-of-sample implied volatilities. However, in the presence of volatility smirk, extrapolation creates a more severe kink, if the current price of the underlying asset is too close to lowest observed strike price. Nonetheless, Jiang and Tian (2005, 1317) argue that in the presence of typically observed smile and smirk patterns, the error caused by extrapolation is still less severe than the error caused by truncation.

5.2 Estimation of the realized volatility

The realized volatility is an abstract concept and its estimates are subject to noise caused by market microstructure factors. As the realized volatility is the reference value in the volatility forecast evaluation, the choice of the estimation procedure has a significant effect on the forecast performance. To begin with, the sample interval of the time series plays an important role in the choice of procedure. If the realized volatility for time interval t to T is estimated, the accuracy of the estimate depends on the amount of information in use: the accuracy increases as the available information inside the interval increases. For example, when the daily data is available and the monthly volatility is required, the realized volatility can be estimated by calculating the sample standard deviation, (Poon 2005, 10)

$$\hat{\sigma} = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}(r_t - \mu)^2}$$

The problem with this approach is that the sample mean is not an accurate estimate for true mean unless the sample period is very long. Since financial assets' return volatility is stochastic and changes over time, it may be convenient to limit the historical data used. While including the older data leads to a larger sample, it may not contain relevant information about the current state of the volatility. Since the volatility is calculated in terms of deviations from the mean, the noisiness of the sample mean has a great effect in the volatility estimate. To increase the accuracy of the volatility estimate, it is usually worthwhile to impose a different estimate of the theoretical mean, and use it instead of the sample mean. (Poon 2005, 10; Figlewski 1997, 82) Imposing a mean of zero and using the root of average squared returns as volatility estimate is a common practice,

$$\hat{\sigma} = \sqrt{\frac{\sum_{t=1}^{T} r_t^2}{T}}. \qquad (5.2)$$

Again, in order to get good estimate for the realized volatility, the information inside the interval between the volatility estimates is important. If the daily realized volatility is required and the available data contains no information about the intraday price fluctuations, the current practice is to use the absolute daily value of return as a proxy for the realized volatility. The use of absolute value as a proxy is equivalent of forcing the time interval T to one and the mean of the return to zero. Even though the absolute value of the daily return is unbiased proxy for the daily volatility, it is extremely noisy. The use of intraday data provides more accurate estimates for the daily volatility. (Poon 2005, 10-11; Andersen & Bollerslev 1998, 886)

In this study, the estimate for the weekly realized volatility is required in order to obtain a reference value for the one-week forecast. Since there is no intraday data available, the weekly realized volatility is estimated as the square root of the weekly average of daily squared returns, where the daily returns are continuously compounded. Since each estimate of weekly volatility is a product of multiple daily returns, it is not likely to be very noisy.

5.3 Forecast Evaluation

When conducting a contest between alternative forecasting models, one of the most important decisions is the choice of evaluation criterion on which the model selection is based. There is no simple way to rank alternative forecasts, since the choice of the evaluation criterion to rank the forecast performance is affected by a number of issues, such as the intended use of the forecast. Regardless of these issues, the discussion on forecast evaluation is typically scarce in the volatility forecasting literature.

To begin with, it is not clear whether the form of predicted variable should be conditional volatility or its squared form, conditional variance. Poon (2005, 21-22) proposes that conditional volatility, σ_t, is the correct form for the predicted variable, and if the conditional volatility cannot be estimated reliably, it is not necessarily worthwhile to perform a comparison between the forecasting models at all. The use of conditional volatility as a predicted variable has several justifications. Firstly, given that the forecasting models are not able to accurately predict shocks that are new to the system, the use of conditional variance, σ_t^2, as a predicted variable gives excessive weight to forecast errors caused by these shocks. Since an excessive weighting is given to the extreme events, the use of conditional variance disfavors the forecast models that tend to produce less extreme forecasts. Secondly, the error statistics favor the choice of conditional volatility as the predicted variable, since the square errors are commonly used to measure forecast error. When the conditional variance is used as a predicted variable, the square error is equal to the fourth power of the same error in terms of conditional volatility. This is problematic because the estimation of the fourth moment of a variable is typically unreliable, which leads to wider confidence intervals. As a result, the differences between competing forecasting models are less likely to be statistically significant and the comparison between them becomes more difficult. (Poon 2005, 22)

5.3.1 Loss functions in forecast evaluation

With the exception of perfect foresight, it is not possible to find forecasting method, which always sets the predicted variable $\hat{X}_t$ equal to the actual outcome X_t. The difference between the predicted variable and the actual outcome is the forecast error, $e_t = \hat{X}_t - X_t$. In order to compare the performance of different volatility forecasts, a formal method of trading off these forecast errors is required. Generally, these errors are evaluated by means of a loss function which gives a mathematical representation of the potential losses caused by the error. In this case, the forecast evaluation is conducted by minimizing the value of the designated loss function. Ideally this loss function would take into account the usefulness of the forecast in economic decision making. Especially, these economic loss functions should incorporate the costs faced by the user of the volatility forecast, which means that the economic problem faced by the forecast user determines the characteristics of the loss function. For example, the optimal loss function for risk-averse investor's portfolio decision problem could be different from the optimal loss function for option market maker. However, the use of economic loss function requires a deep understanding of the forecaster's objectives, such as assumptions about the properties of the utility function, which makes their use difficult. This is why the volatility forecast evaluation is typically conducted by using purely statistical loss function. The statistical loss functions are commonly based on the descriptive statistics of the forecast errors. (Elliot & Timmerman 2008, 12-13; Lopez 2001, 87-88)

The most commonly used statistical loss function in economic forecasting is the mean square error (MSE) which measures the accuracy of a forecast in terms of squared forecast errors. MSE is particularly tractable as a loss function since its interpretation and implementation are straightforward. There are no unknown parameters and the optimal forecast is simply the conditional mean of the predicted variable $\hat{X}_t^* = E(X_t|\Omega_{t-h})$, where Ω_{t-h} is the collection of information available at the forecast origin. The unit of measure for MSE is the squared unit of measure of the predicted variable. Hence the square root of MSE, root mean square error (RMSE), is often used as a loss function instead of MSE, since it has the same unit of measure as the predicted variable. (Elliot & Timmerman 2008, 15; Poon 2005, 23)

MSE and RMSE are defined as,

$$MSE = \frac{1}{N}\sum_{t=1}^{N} \varepsilon_t^2 = \frac{1}{N}\sum_{t=1}^{N}\left(\hat{X}_t - X_t\right)^2,$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N} \varepsilon_t^2} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\hat{X}_t - X_t\right)^2}. \qquad (5.3)$$

Mean absolute error (MAE) is another commonly used statistical loss function which measures the accuracy of the forecast. In a case where MAE is used as a loss function, the optimal forecast is the conditional median of the predicted variable, given that the distribution of the predicted variable is continuous. MAE has the same unit of measure as the predicted variable. Commonly used variant of MAE, mean absolute percent error (MAPE) measures the magnitude of the absolute errors relative to the predicted variable. (Elliot & Timmerman 2008, 15; Poon 2005, 23)

MAE and MAPE are defined as,

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|\varepsilon_t| = \frac{1}{N}\sum_{t=1}^{N}\left|\hat{X}_t - X_t\right|, \qquad (5.4)$$

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\frac{|\varepsilon_t|}{X_t} = \frac{1}{N}\sum_{t=1}^{N}\frac{\left|\hat{X}_t - X_t\right|}{X_t}. \qquad (5.5)$$

It is important to note that these loss functions are themselves subject to error and noise. Therefore, in order to make conclusions on which one of the forecasts is superior, it is necessary to perform tests of statistical significance. Diebold and Mariano (1995) developed three statistical significance tests for testing the equal accuracy between two forecasting models. The three tests relate the forecast errors to some very general loss functions and are designed to analyze the difference between the forecast errors of two competing models. One of the major advantages of these tests of equal accuracy is that they are not dependent on certain loss function. Moreover, simulation studies show that the three tests are robust against a very general class of forecast error distributions (Poon 2005, 25).

Mariano and Diebold (1995) proposed a setting with two forecasts, $\{\hat{X}_{it}\}_{t=1}^{T}$ and $\{\hat{X}_{jt}\}_{t=1}^{T}$, for a predicted time series $\{X_t\}_{t=1}^{T}$. The associated forecast errors are notated as $\{e_{it}\}_{t=1}^{T}$ and $\{e_{jt}\}_{t=1}^{T}$. Moreover, $g(X_t, \hat{X}_{it})$ is the loss function that is used to measure the accuracy of the forecast, which in this case of statistical loss functions presented above (equations 5.3, 5.4 and 5.5) is defined in terms of the forecast errors,

$$g(X_t, \hat{X}_{it}) = g(e_{it}).$$

In the tests introduced by Diebold and Mariano (1995) the equal accuracy of two competing tested by analyzing the differences in loss function values. This loss differential between two competing models is defined as,

$$d_t \equiv g(e_{it}) - g(e_{jt}).$$

Two tests of equal accuracy are used in this study. These are the first two of the three tests proposed by Diebold and Mariano (1995), namely asymptotic test and sign test. The asymptotic test targets on the mean of the forecast loss differential whereas the sign test targets on the median. The test statistic for the asymptotic test is based on the asymptotic distribution for the mean of the loss differentials. In the case of sign test, the sign of the loss differential defines the test statistic. The third test proposed by Diebold and Mariano, sign-rank test, is the one that is not used in this study. The sign-rank test targets on both the median and the rank of the forecast loss differential.

The use of asymptotic test, which targets on the mean of the sample loss differential, can be motivated by examining the expected loss function values at population level. The equal accuracy of two competing forecasts can be defined in terms of expected loss function value: the two forecasts are equally accurate, if their expected loss function values are equal. With this definition, the null hypothesis for equal accuracy can be expressed as $E\big(g(e_{it})\big) = E\Big(g\big(e_{jt}\big)\Big)$. By moving the right-hand side of the equation to the other side, the null hypothesis can be stated as $E(d_t) = 0$. In order to obtain the corresponding null hypothesis at sample level, the expected value of the loss differential is replaced by its sample estimate, that is the sample mean of the loss differentials. Now, the null hypothesis for the Diebold and Mariano (1995) asymptotic test of equal forecast accuracy takes following form,

$$\bar{d} = \frac{1}{T}\sum_{t=1}^{T}\left|g(e_{it}) - g\big(e_{jt}\big)\right| = 0.$$

To test the null hypothesis, the distribution of the sample mean for loss differentials is required. Diebold and Mariano (1995) deduced the asymptotic distribution for the sample mean for loss differentials under the assumptions that the differential series is covariance stationary and of short memory. Based on the asymptotic distribution, the test statistic for the hypothesis is stated as,

$$S_1 = \frac{\bar{d}}{\sqrt{\frac{1}{T}2\pi\hat{f}_d(0)}} \sim N(0,1), \qquad (5.6)$$

where, $\hat{f}_d(0)$ is the spectral density of the loss differential at frequency zero. Following standard practice, Diebold and Mariano (1995) proposed that the weighted sum of sample autocovariances, $\hat{\gamma}_k$, could be used in order to obtain a consistent estimate for $2\pi\hat{f}_d(0)$,

$$2\pi\hat{f}_d(0) = \hat{\gamma}_0 + 2\sum_{k=1}^{T-1}\hat{\gamma}_k.$$

On the other hand, the null hypothesis in the sign test is that the median of the loss differentials is zero,

$$Med(d) = Med\left(g(e_{it}) - g(e_{jt})\right) = 0.$$

The use of sign test, that targets on the median of the loss differentials, can be motivated with its meaningful and intuitive interpretation: under the null hypothesis the two forecasts are equally accurate in the sense that it is equally probable that loss function value for one of the forecast is greater than it is for the other,

$$Pr\left(g(e_{it}) > g(e_{jt})\right) = Pr\left(g(e_{jt}) > g(e_{it})\right).$$

Assuming that the loss differentials are independent and identically distributed, the test statistics is following,

$$S_2 = \sum_{t=1}^{T} I_+(d_t),$$

where,

$$I_+(d_t) = \begin{cases} 1 & if\ d_t > 0 \\ 0 & otherwise \end{cases}.$$

In large sample, the standardized version of the test statistics is asymptotically normal,

$$S_{2a} = \frac{S - 0{,}5T}{\sqrt{0{,}25T}} \sim N(0{,}1). \quad (5.7)$$

These two tests proposed by Diebold and Mariano (1995) can be used in wide variety of forecasting evaluation problems because they are not dependent on a set of restrictive assumptions and the test statistics are valid for a very general class of loss functions. For example, contrary to many other tests of similar purpose, Diebold and Mariano’s tests do not require the loss function to be quadratic. The selected loss function may even be asymmetric and noncontinuous. Moreover, the properties of the forecast error statistics are not restricted as the tests are valid for non-zero mean and non-Gaussian forecast errors. Additionally, the tests are robust for contemporaneously correlated forecast errors. This is particularly important since the errors of two forecasts of the same time series are commonly contemporaneously correlated.

In summary, the loss functions that are defined in terms of forecast errors are the most common way to measure the accuracy of a forecast. However, the values that the loss functions generate are themselves subject to error and noise, which complicates the comparison of loss function values. Diebold and Mariano (1995) proposed a set of tests of equal forecast accuracy which enable statistical inference on the accuracy of competing forecast models. The major advantage of these tests is that they are robust to a very general class of loss functions and error statistics, which is why they are often used in the volatility forecasting literature. The only prerequisite for the selected loss function is that the loss differential, $d \equiv g(e_{it}) - g(e_{jt})$, is defined. If that is the case, the implementation of the tests is relatively straightforward.

5.3.2 Regression-based forecast evaluation

The loss functions that measure the accuracy of the forecast are not able to comprehensively capture the performance of a forecast. This is because the accuracy of the forecast is only one dimension of the forecast performance. Thus loss functions defined in terms of forecast errors are not the only way to evaluate volatility forecasts. For example, a biased forecast that produces higher loss function values may well have more predictive power than an unbiased forecast, if the bias can be corrected. Therefore, it is useful to distinguish between the forecast accuracy and the predictive power of a forecast. Furthermore, it is useful to complement the evaluation criterion by performing tests that are designed to capture the other aspects of the forecast performance in addition to accuracy. McCracken and West (2008, 300) group the forecast evaluation measures into five categories: forecast accuracy, forecast encompassing, forecast efficiency, forecast bias and sign predictability. In addition to the tests of equal accuracy, such as tests proposed by Diebold and Mariano (1995), the economic forecasting literature presents methods for examining the other aspects of the forecast as well. One of the commonly used methods is regression-based forecast evaluation which is used to analyze the forecast encompassing, forecast bias and the predictive power of the forecast.

In order to examine the predictive power and the bias of a volatility forecast, the actual volatility, X_t, is regressed on the volatility forecast, $\hat{X}_t$. The univariate regression – which is also known as Mincer-Zarnowitz regression – takes the following form,

$$X_t = \alpha + \beta \hat{X}_t + v_t. \quad (5.8)$$

This univariate regression plenty of information about the volatility forecast. To begin with, it may be reasonable to examine if the forecast contains any useful information about the actual volatility at all. This can be done by analyzing the regression coefficient β; if $\beta = 0$, the forecast contains no information on the future volatility. Moreover, the R^2 statistic of the regression represents the degree of variation in the actual volatility that is explained by the volatility forecast. Hence the forecast with higher R^2 statistic is considered to contain more predictive power on the actual volatility, than a forecast with lower R^2 statistics. (Andersen & Bondarenko 2007, 14)

The biasedness of the forecast can also be examined with regression-based methods. In the univariate regression (equation 5.8), an unbiased forecast is subject to constraint on the regression coefficients that $\alpha = 0$ and $\beta = 1$. The forecast is upwardly (downwardly) biased if one or both of the coefficients α and β is smaller (greater) than zero. In a case where $\alpha > 0$ and $\beta < 1$, the forecast tends to be too low in times of low volatility and too high in times of high volatility. This is probably the most common result in volatility forecasting literature, since the models are not designed to capture shocks that are new to the system, and thus are not able to forecast most extreme volatility observations. (Poon 2005, 118)

Regression-based forecast evaluation is also advantageous when comparing alternative forecasting models and procedures. Regression-based methods can be used to analyze the forecast encompassing, which is closely related to the literature of forecast combination. The idea of forecast combination is to use more than one model to generate a volatility forecast and by doing so make use of richer set of information compared to the separate forecasts. Hence, the forecast generated by combining two or more forecasts should perform better than those forecasts would perform separately. However, this is not always the case, since the combined forecast does not necessarily include incremental information on the actual volatility compared to the information content of the best individual forecast. In this case the forecast combination is unnecessary since the best forecast encompasses the information that the other forecasts contain. To examine the information content of two competing forecasts, another forecast, $\hat{X}_{2,t}$, is added to the univariate regression (equation 5.8). The encompassing regression takes the following form,

$$X_t = \alpha + \beta_1 \hat{X}_{1,t} + \beta_2 \hat{X}_{2,t} + u_t. \qquad (5.9)$$

The forecast encompassing can be defined based upon this regression. A forecast generated by model 1, $\hat{X}_{1,t}$, is said to encompass the forecast generated by model 2, $\hat{X}_{2,t}$, if the regression adjusted regression R^2 does not increase significantly when the second forecast is added to the regression. To put in other words, the first forecast is said to subsume the information content of the second forecast. Another way to analyze if the first forecast encompasses the second is to test the regression coefficients: if the regression coefficients meet the condition $\beta_1 > 0$ and $\beta_2 = 0$, the first forecast encompasses the second one. Intuitively, the second forecast contains no incremental information on the future volatility after the information content of the first forecast is used.

There is also a third way to test the information content of the forecasts by analyzing to regression residuals. This can be done by conducting an orthogonality test by regressing the residual of the univariate regression (equation 5.8) against the second forecast, (McCracken & West 2008, 302)

$$v_t = \gamma \hat{X}_{2,t} + e_t$$

If regression coefficient for the second forecast, γ, is not different from zero, when the residual of the univariate regression (equation 5.8) is regressed on the second forecast, the first forecast is said to subsume the second one. If the regression coefficient is different from zero, the information in the second forecast has some value in anticipating the residual of the univariate regression, which implies that the second forecast contains incremental information on the future volatility. However, this test is subject to some serious methodological issues, which is why it is not used in this study (see, for example Newbold & Harvey 2008, 274-275).

In summary, both error statistics and regression-based methods are useful when the forecast performance of alternative models is compared. In order to gain a comprehensive view on the forecasting performance of the selected models, both methods are also applied in this study. The loss function values are analyzed with the above presented tests of equal forecast accuracy by Diebold and Mariano to gain an understanding on the significance on the differences in loss function values. Moreover, in order to gain insight on the predictive power, forecast bias and forecast encompassing, relevant univariate and encompassing regressions are estimated, and the statistical tests are performed on the regression coefficients. The actual analysis on the forecast performance can be found in the section 6.3 of the next chapter.

5.4 DATA

The option data used in this study consists of Friday closing prices across the quoted strike price range for DAX index options. DAX (*Deutschen Aktien IndeX*) is a German stock market index which performance is calculated as a total return of the stocks of 30 largest companies in Frankfurt Stock Exchange. The company selection for DAX index is based on order book volume and the market capitalization. The weighting of the index is based on the market capitalization as well. DAX is a total return index which means that the cash distributions are assumed to be reinvested back into the index. Hence, the estimation of dividend payments is unnecessary. Moreover, the index is adjusted for stock splits and changes in capital, which further simplifies the treatment of data. The data sample for the underlying asset, DAX index itself, consists of daily closing prices from November 16, 2001 to September 14, 2012. 1-month Euribor rate is used as a proxy for the risk-free rate. All of the data used in this study was gathered from Thomson Reuters Datastream.

The DAX index options, which are traded in EUREX derivatives exchange, are among the most liquid options in the world. In addition, the range of quoted strike prices for DAX index options is wide. Due to these facts, the option prices should contain relatively small amount of noise caused by market microstructure factors. The options are of European type, which means they can only be exercised at expiry. Hence, the issues related to early exercise are avoided in this study. The sample period for weekly observations of call option prices is from September 8, 2006 to September 7, 2012. The weekly observations are the closing prices for the call options from Friday. Following Jiang and Tian (2005), options with one week or less to expiration are excluded from the sample, since their prices are subject to market microstructure noise. Besides that, the options that are closest to expiration are used to determine the implied volatility measures.

Weekly realized volatility, which is the forecasted variable, is estimated by calculating the square root of the weekly average of daily squared returns. The three volatility forecasts – model-free implied volatility, Black-Scholes implied volatility and GARCH(1,1) – are estimated from September 8, 2006 to September 7, 2012. The forecast data set for each of the forecasting methods consists of 314 one-week horizon forecasts. The implied volatility forecasts are estimated by using non-overlapping sample periods and hence the telescoping problem in option price data is avoided (see Christensen, Hansen, Prabhala 2001).

The model-free implied volatility forecast is generated by applying numerical integration to equation 5.1. In order to avoid problems caused by thin trading and other market microstructure factors, the options with moneyness level outside the interval 0.96 to 1.04 are excluded from the sample. The moneyness level is defined as X/S where X is the strike price of the option and S is spot price of the underlying asset (cf. Jiang & Tian 2005). Additionally, the options violating the arbitrage boundaries are excluded from the sample.

As the first part of the actual implementation of the model-free implied volatility, the truncated set of call option prices is converted to Black-Scholes implied volatilities (BSIV). Then, cubic spline interpolation is applied to the set of implied volatilities in order to create a grid of strike price-BSIV observations inside the defined moneyness interval. After the interpolation, the grid contains a BSIV value for strike prices with frequency of 0.1 index points, inside the moneyness interval defined above. The implied volatilities outside the interval are approximated as the values at the end-points of the interval, which forces the tails to be lognormal. The next step is to convert BSIV's back to call option prices by using Black-Scholes model (equation 4.1). It is important to note that the use of Black-Scholes pricing equation is only used as a computational tool, and its use in this manner does not require its assumptions to hold. Finally, the numerical integration is done by applying trapezoidal rule to equation 4.1 to get the model-free implied volatility. The Black-Scholes implied volatility series is calculated from the nearest-to-the-money call option price by using the equation 5.1. Nearest-to-the-money option is the one with strike price nearest to the spot price of the underlying asset.

The GARCH(1,1) model is used as a competing forecast model for the implied volatility measures. GARCH(1,1) is often found to outperform more sophisticated time series models which is why it is commonly used as benchmark in model comparison (see, for example Ederington & Guan 2005). In this study, the GARCH(1,1) forecasts are generated with rolling scheme. This means that for each time period, an out of sample forecast is generated, by estimating a GARCH(1,1) model using the DAX index returns calculated from 250 latest weekly price observations. This means that the GARCH(1,1) forecast for period $n+1$ is based on information set $\{n-250,\cdots,n\}$. The GARCH(1,1) model is estimated via maximum likelihood procedures and is defined by the following equation,

$$r_t = \mu_t + \varepsilon_t,$$

$$\varepsilon_t = \sigma_t z_t,$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \sigma_{t-1}^2 + \alpha_2 \varepsilon_{t-1}^2.$$

Where,
r_t is the logarithmic return on the financial asset,
μ_t is the conditional mean of the return process,
σ_t is the conditional volatility of the return process, and,
$z_t \sim N(0,1)$ is a white noise process.

The characteristics of the resulting realized volatility and volatility forecast series are analyzed in the section 6.1. Additionally, the properties of the time series for the DAX index are analyzed. In the last section of this chapter, 6.2, the forecast performance of these volatility measures is evaluated by using forecast errors and some regression-based applications. A more detailed view on the procedures of forecast performance evaluation is provided in section 5.3.

6 EMPIRICAL RESULTS

This chapter presents the empirical findings of the study. The first section of the chapter provides some descriptive statistics for DAX index and for the volatility series: model-free implied volatility, Black-Scholes implied volatility and GARCH(1,1) forecast. Moreover, the second section of this chapter contains the results of the analysis on the forecast performance of these three volatility series. This analysis includes the examination of the accuracy of the forecasts by performing tests on the forecast error statistics. Additionally, the information content and the biasedness of the forecasts are examined by performing regression-based tests on the forecasts.

6.1 Descriptive statistics of the variables

This section provides some descriptive statistics for DAX index and the estimated volatility series. In addition to the examination of the statistical properties of these variables, the macroeconomic factors that have driven the DAX index value during the sample period are briefly discussed. Furthermore, some graphical presentation of DAX index value and the model-free implied volatility is provided.

Figure 6.1 depicts the DAX index value on weekly a basis from September 15, 2006 to September 14, 2012. The weekly observations consist of the Friday closing prices. As the graph explicitly shows, the value of DAX index is significantly affected by the turbulent and highly volatile times in the global capital markets. Both the effects of the financial crisis of 2007-2008 and the contributing European sovereign-debt crisis can be seen in the index value. The highest point in DAX index value inside the sample interval was reached at the eve of the financial crisis on July 13, 2007, and the lowest index value was observed in the aftermath of the crisis on March 6, 2009. Another significant plunge was observed in the first half of 2011, in the midst of the European sovereign-debt crisis. In conclusion, the sample period consists of times of high volatility in financial market, which may affect the forecasting performance of model-free implied volatility and the other forecasts.

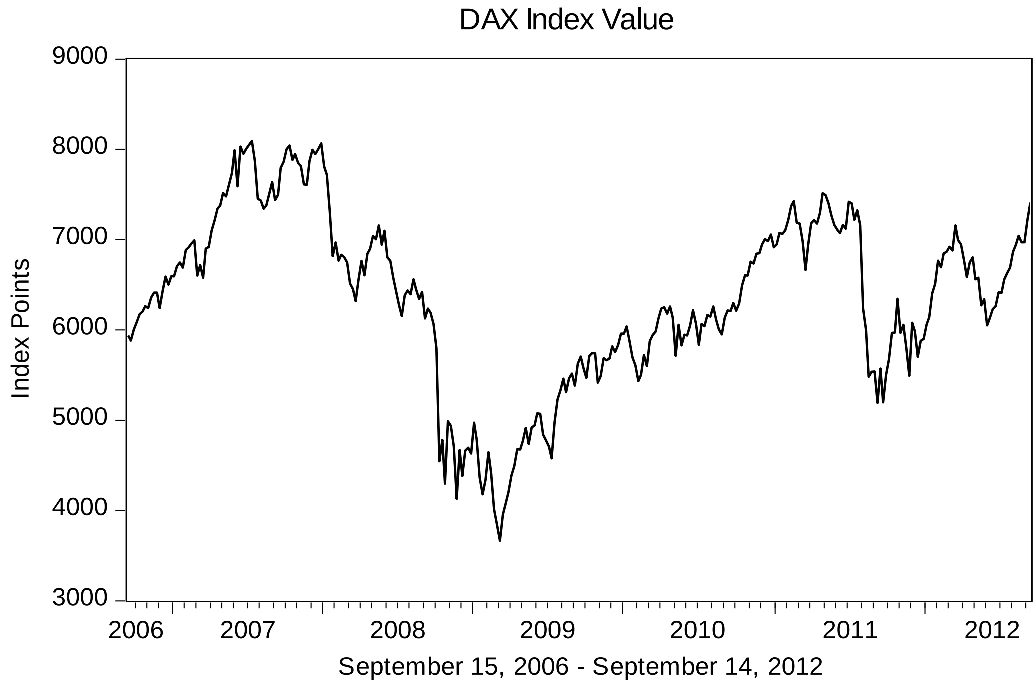

Figure 6.1 The value of DAX index

Table 6.1 provides the descriptive statistics for the DAX index value and for its weekly returns in the sample period. The weekly return is defined as the first logarithmic difference of the weekly index values multiplied by 100. Consistent with the earlier empirical work, the observed weekly return distribution for the index is left-skewed and platykurtic. Moreover, while the mean of the weekly returns in the sample period is close to zero, the weekly median return is at 0.51 which makes 26.64 on annual basis. The minimum weekly return in the sample period, -24.3, was observed on October 10, 2008 in the midst of the financial crisis. The highest weekly return, 14.9, was observed three weeks later on October 31, 2008.

Table 6.2 provides descriptive statistics for the model-free implied volatility (MFIV), Black-Scholes implied volatility (BSIV), GARCH(1,1) forecast and the realized volatility (RV). All forecasts are estimated weekly with forecasting horizon of one-week, from Friday to Friday. The forecasts are generated for the time period September 15, 2006 to September 14, 2012 which results in sample size of 314.

Table 6.1 Descriptive statistics of DAX index

TABLE 6.1
Descriptive statistics of DAX Index, $n = 314$

	Index Value	**Weekly return**
Mean	6335.60	0.0709
Median	6407.25	0.509
Min	3666.41	−24.347
Max	8092.77	14.942
Standard Deviation	959.26	3.682
Skewness	-0.439	−1.094
Kurtosis	2.766	10.501

Notes: The descriptive statistics are for weekly observations from Friday 15/9/2012 to Friday 14/9/2012. Weekly return is defined as logarithmic return multiplied by 100.

To begin with, according to the means and the medians of the volatility series in Table 6.2, all three forecasts seem to be positively biased. The same descriptive statistics indicate that model-free implied volatility is more biased forecast for the realized volatility than the two alternative forecasts. Moreover, the standard deviation of the realized volatility measure seems to be higher than the standard deviations of the forecasts in the sample period, and all observations of two implied volatility forecasts fall inside the range of minimum and maximum realized volatilities. However, this is not the case with GARCH(1,1) forecast as its maximum value is higher than the maximum value of the realized volatility. All volatility series are right-skewed and platykurtic.

Table 6.2 Descriptive statistics for the volatility series

TABLE 6.2
Descriptive statistics for the volatility series, $n = 314$

	RV	**MFIV**	**BSIV**	**GARCH(1,1)**
Mean	0.216	0.253	0.235	0.234
Median	0.183	0.231	0.210	0.186
Min	0.036	0.113	0.109	0.121
Max	1.038	0.735	0.704	1.402
Std. Dev.	0.142	0.102	0.099	0.139
Skewness	2.190	1.742	1.900	3.831
Kurtosis	9.620	6.861	7.617	24.028

Notes: The sample period is from Friday 15/9/2012 to Friday 14/9/2012.

To begin with, according to the means and the medians of the volatility series in Table 6.2, all three forecasts seem to be positively biased. The same descriptive statistics indicate that model-free implied volatility is more biased forecast for the realized volatility than the two alternative forecasts. Moreover, the standard deviation of the realized

volatility measure seems to be higher than the standard deviations of the forecasts in the sample period, and all observations of two implied volatility forecasts fall inside the range of minimum and maximum realized volatilities. However, this is not the case with GARCH(1,1) forecast as its maximum value is higher than the maximum value of the realized volatility. All volatility series are right-skewed and platykurtic.

Figure 6.2 depicts model-free implied volatility (MFIV) series and the realized volatilities (RV) in the sample period. The figure shows that the MFIV forecast generally moves together with the estimated realized volatility series. However, as the descriptive statistics suggest, the model-free implied volatility seems to be positively biased forecast for the realized volatility and it is not able to forecast the most extreme volatility observations. Moreover, in line with the earlier empirical findings, volatility's tendency to cluster is evident in the figure. Even though model-free implied volatility tends to produce less extreme forecasts compared to the realized volatility in times of high volatility, it seems that it is still able to predict the shocks in volatility to some extent. This is with one exception; model-free implied volatility was not able to predict the peak in volatility in January 2008, when the fears of upcoming recession lead to a sudden crash in share prices globally, which caused a peak in the realized volatility.

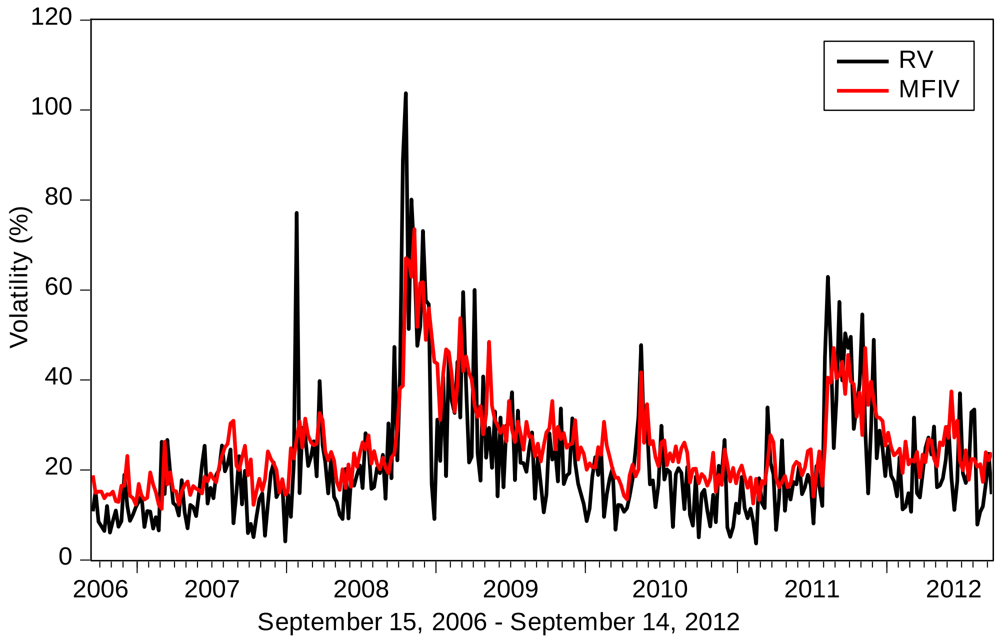

Figure 6.2 Model-free implied volatility and the realized volatility

6.2 The results of the forecast performance analysis

This section provides the analysis on the forecast performance of model-free implied volatility. The methodology is based on the forecast evaluation criterion discussed in section 5.3. This section begins with a comparison of loss function values that are defined in terms of error statistics. The loss function values generated by the model-free implied volatility are compared to the error statistics of forecasts generated with two alternative models, namely Black-Scholes implied volatility and GARCH(1,1). All of the three loss functions used are purely statistical representations of the forecast errors. The three loss functions applied in this study are: root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE). The loss functions are defined by equations 5.3, 5.4 and 5.5. Values for these three loss functions are computed for each of the forecasts, and the results are analyzed using the tests of equal accuracy proposed by Diebold and Mariano (1995).

Table 6.3 provides the error statistics for model-free implied volatility, Black-Scholes implied volatility and GARCH(1,1) forecast. Black-Scholes implied volatility forecast generates the lowest loss function values for all of the three functions. In the case of root mean square error, the difference in loss function value between the Black-Scholes implied volatility and model-free implied volatility is statistically significant according to the asymptotic and the sign tests of Diebold and Mariano (1995). Moreover, the model-free implied volatility yields significantly lower loss function value than GARCH(1,1) forecast, but only according to the asymptotic test.

Table 6.3 Error statistics for the volatility forecasts

TABLE 6.3
Forecast Error Statistics, $n = 314$

Model	RMSE		MAE		MAPE	
	Level	Rank	Level	Rank	Level	Rank
MFIV	0.1039	2	0.0780	2	0,5113	3
BSIV	0.0953	1^{***}_{***}	0.0681	1^{***}_{***}	0.4274	1^{***}_{***}
GARCH	0.1210	3^{***}	0.0845	3^{***}	0,5089	2^{***}

Notes: Error statistics for one-week horizon forecasts from 15/9/2006 to 14/9/2012. *, **, *** indicate significant difference from the MFIV loss function value at the levels 0.05, 0.01, 0.001. In the superscript, asterisks denote significant difference from MFIV loss function value according to Diebold and Mariano’s (1995) asymptotic test. In the subscript, asterisks denote significant difference from MFIV loss function value according to Diebold and Mariano’s sign test.

In the case of mean absolute error, the forecasts are ranked similarly and the tests of statistical significance yield identical results with the case of root mean square error. However, in the case of mean absolute percent error the results differ from the ones of the RMSE and MAE: now, model-free implied volatility yields a higher loss function value than GARCH(1,1) forecast. Again, the difference in loss function values for Black-Scholes implied volatility and model-free implied volatility is statistically significant for mean absolute percentage errors. Moreover, the difference in MAPE loss function values between model-free implied volatility and GARCH(1,1) is statistically significant according to asymptotic test, but the null hypotheses of equal accuracy cannot be rejected for the sign test.

To summarize the results on forecast accuracy, model-free implied volatility yields significantly higher forecast errors than Black-Scholes implied volatility, when the most common error statistics are applied. This indicates that the forecast accuracy of Black-Scholes implied volatility is superior to the forecast accuracy of model-free implied volatility. Moreover, according to the error statistics, model-free implied volatility forecasts realized volatility more accurately than GARCH(1,1) model, but the difference in accuracy is not as unambiguous.

While the first half of this section focused on the analysis of forecast accuracy for the model-free implied volatility and the competing models, the latter part of the section focuses on the other dimensions of the forecast performance. The predictive power, forecast bias and the information content of the forecasts are examined with regression-based forecast evaluation criterion. The predictive power and the biasedness of the forecasts are measured by testing the coefficients in a regression where predicted variable, namely realized volatility, is regressed on the predictive variable (see equation 5.8). In addition to the coefficient tests, the regression R^2 statistic is used as a measure of predictive power. The rationale for this is that the R^2 statistic measures the degree of variation in predicted variable explained by the predictive variable. Furthermore, the regression-based methods are used to examine the information content of the forecasts. The analysis is done with univariate and encompassing regressions (see equation 5.9), where the regression adjusted R^2 statistics are used to determine the information content of the forecasts. Additionally, the information content is examined by performing tests on the estimated coefficients of the encompassing regressions.

In the univariate regressions, ex post realized volatility is simply regressed against the volatility forecast. Let the ex post realized volatility for week $t+1$ to be denoted by σ_{t+1} and the volatility forecast j among a set of J forecasts by $\hat{\sigma}_{j,t}$, $J = MFIV, BSIV, GARCH$. Now, the univariate regressions take the form,

$$\sigma_{t+1} = \alpha_j + \beta_j \hat{\sigma}_{j,t} + u_{j,t+1}.$$

This univariate regression is used to evaluate the predictive ability of the forecast $\hat{\sigma}_j$. If the regression coefficient satisfies the condition $\beta_j \neq 0$, the forecast $\hat{\sigma}_j$ contains useful information on the ex post realized volatility σ. Moreover, for an unbiased forecast $\hat{\sigma}_j$ the regression coefficients are subject to constraint that $\alpha_j = 0$ and $\beta_j = 1$. Additionally, the regression R^2 statistic is used to measure the predictive power of the forecast.

In the encompassing regressions, two or all three of the forecasts are regressed together against the ex post realized volatility. The encompassing regressions take the following form,

$$\sigma_{t+1} = \alpha + \beta_{MFIV}\hat{\sigma}_{MFIV,t} + \beta_{BSIV}\hat{\sigma}_{BSIV,t} + u_{t+1},$$

$$\sigma_{t+1} = \alpha + \beta_{MFIV}\hat{\sigma}_{MFIV,t} + \beta_{GARCH}\hat{\sigma}_{GARCH,t} + u_{t+1},$$

$$\sigma_{t+1} = \alpha + \beta_{BSIV}\hat{\sigma}_{BSIV,t} + \beta_{GARCH}\hat{\sigma}_{GARCH,t} + u_{t+1},$$

$$\sigma_{t+1} = \alpha + \beta_{MFIV}\hat{\sigma}_{MFIV,t} + \beta_{BSIV}\hat{\sigma}_{BSIV,t} + \beta_{GARCH}\hat{\sigma}_{GARCH,t} + u_{t+1}.$$

The information content of the forecasts is analyzed with the encompassing regressions. If the adjusted R_2 statistic does not increase, when an additional forecast is added to the regression, the added forecast contains no incremental information on the ex post realized volatility. The information content of the forecasts is also examined by performing tests on the regression coefficients. If the regression coefficient for one of the forecasts differs from zero, while it is zero for the other, the first forecast encompasses the second one. Moreover, the tests on regression coefficients are used to analyze if the leading forecast is unbiased and encompasses the other forecasts in the regression. In all but one of the regression, the leading forecast is model-free implied volatility. The exception is the encompassing regression where only Black-Scholes implied volatility and GARCH(1,1) forecast are regressed against ex post realized volatility. In this case, Black-Scholes implied volatility is the leading forecast.

Table 6.4 provides the results of the univariate and encompassing regressions. To begin with, the slope coefficients from univariate regressions (1-3) are positive and statistically significant for all three forecasts. This result is consistent with the earlier research and indicates that the two implied volatility measures and GARCH(1,1) forecast are all valid forecasts as they contain information on the future realized volatility. According to the adjusted R^2 statistics of the univariate regressions, Black-Scholes implied volatility has the best predictive power, as it explains 57.0 % of the variation in future realized volatility, whereas the adjusted R^2 statistics for the model-free implied volatility and GARCH(1,1) forecast are 53.1 % and 40.5 % respectively.

Table 6.4 Univariate and encompassing regressions

TABLE 6.4
Univariate and encompassing regressions for next-week realized volatility (OLS)

Regression	α	β_{MFIV}	β_{BSIV}	β_{GARCH}	$adj. R^2$	F stat
(1) MFIV	-0.040 (-2.037)	1.013 (11.075)			0.531	33.491 (0.000)
(2) BSIV	-0.037 (-2.211)		1.078 (12.780)		0.570	13.147 (0.000)
(3) GARCH	0.063 (4.403)			0.651*** (11.160)	0.405	18.285 (0.000)
(4) MFIV + BSIV	-0.026 (-1.781)	-0.735 (-2.184)	1.823 (4.848)		0.574	13.836 (0.000)
(5) MFIV + GARCH	-0.035 (-2.144)	0.800 (8.019)		0.209 (3.051)	0.546	4.658 (0.010)
(6) BSIV + GARCH	-0.0347 (-2.281)		0.935 (8.759)	0.132 (1.895)	0.573	2.302 (0.102)
(7) MFIV+BSIV+GARCH	-0.025 (-1.785)	-0.669 (-2.154)	1.629 (4.589)	0.117 (1.710)	0.578	11.074 (0.000)

Notes: The heteroskedasticity and autocorrelation consistent t-statistics are reported in parentheses below the parameter estimates for regression coefficients (Newey & West 1987). For univariate regressions 1-3 *, ** and *** indicate significant difference from one for the forecast coefficient β_i and the F test is for the joint hypothesis H_0: $\alpha = 0$ and $\beta_i = 1$ ($i = MFIV, BSIV, GARCH$). For multivariate regressions 4-5 the F test is for the hypothesis H_0: $\beta_{MFIV} = 1$ and $\beta_j = 0$ ($j = BSIV, GARCH$). For multivariate regression 6 the F test is for the hypothesis H_0: $\beta_{BSIV} = 1$ and $\beta_{GARCH} = 0$, and for regression 7 the F test is for the hypothesis H_0: $\beta_{MFIV} = 1$ and $\beta_{BSIV} + \beta_{GARCH} = 0$. P-values for the F tests are reported in parentheses below the test statistic values.

The Wald F-tests in the univariate regressions indicate that all three forecasts are biased. Even though the null hypothesis H_0: $\beta_{MFIV} = 1$ is not rejected in regression 1, the Wald F-test rejects the unbiasedness model-free implied volatility forecast. The estimated regression indicates that $\alpha < 0$ and $\beta_{MFIV} = 1$, which implies that model-free implied volatility is positively biased forecast and the bias is relatively stable regardless of the level of volatility. This result is consistent with the theory, which states that a volatility forecast made under risk-neutral probabilities is likely to be upwardly biased in the presence of variance risk premium. In the case of Black-Scholes implied volatility (regression 2), the results of the regression coefficient tests are similar: the forecast is upwardly biased, and the bias is not sensitive to the level of volatility. This is again consistent with the theory as Black-Scholes model assumes constant volatility and hence the pricing equation does not incorporate the variance risk premium. In the case of GARCH(1,1) forecast, the Wald F-test for unbiasedness is rejected as well, and the estimated regression coefficients indicate that $\alpha > 0$ and $\beta_{GARCH} < 1$. These coefficient values imply that the GARCH(1,1) underforecasts low volatility and overforecasts high volatility.

The results from the estimated encompassing regressions are consistent with the conclusions drawn from the error statistics and univariate regressions: Black-Scholes implied volatility's forecast performance is superior to the forecast performance of model-free implied volatility or GARCH(1,1) forecast. The adjusted R^2 statistics indicate that even the combination of model-free implied volatility and GARCH(1,1) forecast seems to have lower predictive power than Black-Scholes implied volatility alone. The adjusted R^2 statistic of the encompassing regression with model-free implied volatility and GARCH(1,1) forecast is 54.6 %, 2.4 percentage points lower than adjusted R^2 of the regression 2. On the other hand, GARCH(1,1) seems to contain only little incremental information on the future volatility above the model-free implied volatility. Even though the regression coefficient for GARCH(1,1) term is statistically significant at the level 0.01, the adjusted R^2 statistic for the encompassing regression increases by only 1.5 percentage points compared to the univariate regression with model-free implied volatility alone.

The results indicate that Black-Scholes implied volatility subsumes nearly all of the information contained by the other forecasts. When the model-free implied volatility and GARCH(1,1) forecast are added to the univariate regression with Black-Scholes implied volatility, there are only slight increases of 0.4 and 0.3 percentage points in the adjusted R^2 statistic. The combination of all three forecasts increases the amount of explained variation in realized volatility by 0.8 percentage points compared to the regression with only Black-Scholes implied volatility as a predictive variable. Additionally, in these two regressions the coefficient for the GARCH(1,1) term is not statistically significant at the level 0.05. In the Wald F-test for regression 6, the null hypothesis $\beta_{BSIV} = 1$ and $\beta_{GARCH} = 0$ cannot be rejected, which supports the hypothesis that the Black-Scholes implied volatility subsumes the information content of GARCH(1,1) forecast. In the case of information content of Black-Scholes implied volatility against the information content of model-free implied volatility, the result is not as clear. When both implied volatility measures are regressed against realized volatility, the coefficient β_{MFIV} is statistically significant at the level 0.05, but not at level 0.01.

The results provided in this chapter enable the analysis of the hypotheses presented in section 1.2. The first hypothesis stating that model-free implied volatility contains information about the future volatility is clearly supported by the results. The test performed on the data, indicate that the second hypothesis, which states that model-free implied volatility is an upwardly biased estimate for the future volatility, is also supported by the data. However, the third and fourth hypotheses are rejected based on the tests performed for the data. In the case of the third hypothesis, which states that model-free implied volatility is more accurate forecast than Black-Scholes implied volatility and GARCH(1,1) forecast, the hypothesis is explicitly rejected. Although the results suggest that model-free implied volatility is more accurate than GARCH(1,1) forecast, it is significantly less accurate than Black-Scholes implied volatility. The fourth hypothesis is similarly rejected. The results do not support the hypothesis that model-free implied volatility subsumes the information content of Black-Scholes implied volatility and GARCH(1,1) forecast. Instead, the results indicate that Black-Scholes implied volatility subsumes nearly all of the information contained by model-free implied volatility.

In summary, the results of this study indicate that the Black-Scholes implied volatility is a better forecast for the future realized volatility than model-free implied volatility, and both implied volatility measures perform better than GARCH(1,1) forecast. The evaluated aspects of the forecast performance include forecast accuracy, forecast biasedness, predictive power of the forecast, and the information content of the forecast. The results on forecast accuracy, predictive power and the information content are almost unambiguous: Black-Scholes implied volatility performs significantly better than its competitors in every measured area of forecast evaluation. The difference in forecasting ability between model-free implied volatility and GARCH(1,1) forecast is not as explicit, but the majority of tests indicate that the difference is significant. The results on the biasedness of the forecasts are consistent with the theory: both implied volatility measures contain systematic upward bias. An interesting property of this bias is that it seems to be stable regardless of the level of volatility. Furthermore, the results indicate that Black-Scholes implied volatility subsumes nearly all of the information contained by model-free implied volatility and GARCH(1,1), since the predictive power of combined forecast with all three forecasts is only slightly better than the predictive power of Black-Scholes implied volatility alone. The combination of model-free implied volatility and GARCH(1,1) leads to a slight increase in predictive power compared to the model-free implied volatility alone. This implies that model-free implied volatility does not subsume all of the information in GARCH(1,1) forecast.

7 SUMMARY AND CONCLUSIONS

This chapter presents the conclusions drawn from the results of this study. The chapter begins with a brief summary on the contents of this study. After the revision on the contents and the objectives, the results presented in previous chapter are compared to the results of prior research on the subject. Besides that, the results are compared to theoretical work on the subject. The chapter ends with a discussion on issues related to the validity and reliability of this study.

The objective of this study is to provide more information on the forecast performance of a more recent volatility measure, model-free implied volatility. The forecast performance is examined by comparing it to two widely used benchmark forecasts. The benchmark forecasts are Black-Scholes implied volatility and forecast series produced with GARCH(1,1) model. The two implied volatility measures, model-free implied volatility and Black-Scholes implied volatility, belong to the class of volatility forecasts that are extracted from the option prices, whereas GARCH is based on time series modeling of conditional volatility. The comparison of the forecast performance of Black-Scholes implied volatility and the time series models has been widely researched topic, but the forecast performance of the more recent implied volatility measure – model-free implied volatility – is still relatively unclear. In this study, the actual comparison of the forecasts is conducted by testing four hypotheses that capture various aspects of forecast performance. These aspects are related to the information content and to the accuracy of the forecasts.

Model-free implied volatility is a theoretical construct which represents the markets' expectations for the future volatility. The expectations are extracted from the prices in the option market. Model-free implied volatility builds on theoretical work on derivatives pricing that goes back to the concept of arbitrage-free markets. The absence of arbitrage enables the modeling of the financial market, which in turn enables the pricing of options and other contingent claims. The prevailing pricing method for all contingent claims that builds on the concept of no-arbitrage is referred to as risk-neutral valuation. The most important insight of this pricing method is that the risk that is specific to a certain contingent claim can be completely hedged. Therefore, the fair value for any contingent claim is simply its expected payoff discounted to present value at risk-free rate.

The development of option pricing theory, especially the invention of risk-neutral valuation, made the research on option implied return distributions possible. Breeden and Litzenberger (1978) were among the first to study this subject, and they presented a way to extract risk-neutral density from the option prices. Risk-neutral density is the risk-neutral return distribution for the underlying asset, implied by the option prices. Hence it can be seen as the market expectation of the return distribution shifted into risk-neutral world. The findings of Breeden and Litzenberger are particularly interesting, since the derivation of model-free implied volatility, first presented by Britten-Jones and Neuberger (2000), begins with the risk-neutral density function of Breeden and Litzenberger. Model-free implied volatility is effectively the amount that the markets are willing to pay in order to receive the sum of squared returns between current time and an arbitrary date in future.

In theory, model-free implied volatility is more advanced construct than the commonly used Black-Scholes implied volatility, and hence its forecast performance should be superior. The most important theoretical advantages of model-free implied volatility are its independency of any specific option pricing model, and the fact that it makes use of information in option prices across the strike price range. On the other hand, model-free implied volatility is derived from the risk-neutral density, which means that it is not an unbiased forecast for the future volatility. Because model-free implied volatility is a risk-neutral forecast, it does not incorporate the premium associated with volatility risk. Academic research on the volatility risk suggests that it is priced with a negative premium which indicates that model-free implied volatility is upwardly biased forecast. However, prior research indicates that the bias caused by volatility risk premium is negligible, and, hence, model-free implied volatility is likely to be relevant volatility forecast in the real world.

The implementation of the model-free implied volatility is subject to several issues. Among these, are the issues caused by numerical integration and the limited availability of strike prices. The methods that are used in this study to solve these issues follow closely the methodology of Jiang and Tian (2005). Another important methodological issue is the estimation of the latent realized volatility factor. In this study the weekly realized volatility is estimated as the square root of the weekly average of daily squared returns. Third important methodological issue in this study is how to evaluate the three competing forecasts. Since the objective of this study is to measure different aspects of forecast performance, both loss function and regression-based forecast evaluation criterion are used. The loss functions are used to measure the accuracy of the competing forecasts, whereas the regression-based methods are used to measure the bias and the information content of the forecasts.

The forecasted variable in this study is a German stock market index DAX. The options on DAX index are among the most liquid ones in the world which means that the market conditions are as close as possible in the real world to the conditions assumed by the implied volatility measures. The different aspects of the forecast performance of model-free implied volatility are evaluated by comparing its performance against two commonly used volatility forecasting models, namely Black-Scholes implied volatility and GARCH. Besides their widespread use, the competing models are chosen due to their good performance in the earlier comparisons of the forecasting ability. Generally, Black-Scholes implied volatility is considered to be superior to forecasts generated with time series models, such as GARCH model family. Even though there are more sophisticated time series models than GARCH models, a simple GARCH(1,1) model is often found to outperform the more sophisticated models, which is why it is chosen as a benchmark forecast in this study.

The results of this study strongly suggest that model-free implied volatility contains useful information on the future ex post realized volatility. In addition to that, the results also suggest that Black-Scholes implied volatility is superior volatility forecast compared to model-free implied volatility. This result applies for every measured aspect of forecast performance. Moreover, both implied volatility measures are found to outperform the forecasts generated with GARCH(1,1) model.

The prior research on the forecast performance of model-free implied volatility is relatively scarce. Jiang and Tian (2005) were the first to implement model-free implied volatility in practice, and they found that it outperforms various other volatility forecasts, such as Black-Scholes implied volatility and a number of time series forecasts, including naïve and more sophisticated ones. Yet, the results of Andersen and Bondarenko (2007) and Cheng and Fung (2012) are consistent with the results of this study. They found that Black-Scholes implied volatility contains more information on the future ex post realized volatility than model-free implied volatility. Like this study, these three studies are all conducted for stock market indices. Taylor, Yadav and Zhang (2010) conducted a similar forecast performance comparison for individual stocks. They found that model-free implied volatility outperforms Black-Scholes implied volatility and times series forecasts in roughly one third of the cases. Consistent with the results of this study, they found that in most cases, Black-Scholes implied volatility performs the best.

The results of this study are also consistent with the theory which suggests that model-free implied volatility is upwardly biased estimate for the ex post realized volatility. However, contrary to the theoretical view, Black-Scholes implied volatility performs better than model-free implied volatility in volatility forecasting. There are a number of possible reasons for that: lack of trading for in-the-money and out-of-the-money options, the issues regarding the volatility risk premiums and the issues that are present when model-free implied volatility is implemented into practice. In conclusion, according to this study and most of the other studies on the subject, model-free implied volatility does not perform as well as Black-Scholes implied volatility in volatility forecasting. Unlike the prior research, this study is not limited to examine the information content of the forecasts, but includes analysis on the forecast accuracy as well.

The forecast performance of the selected forecasts is only examined on a one-week horizon, and the analysis is limited to time period ranging from 2006 to 2012, which certainly were not typical times in terms of financial market volatility. Another thing to consider is that DAX index options are among the most liquid options in the world, and the performance of implied volatility measures might be significantly poorer with less liquid options. Moreover, the choice of realized volatility measure may have a great effect on the results of the study. Due to these factors, the results of this study are not necessarily robust for different forecasting horizons, option markets and realized volatility measures. Yet, most of the prior studies on this subject have reached results that are consistent with the results of this study. In addition, this study complements the earlier research – that concentrates on the information content – by examining the forecast accuracy as well. The methodological choices of the study are consistent with the prior research, and the tests performed to the data are repeatable.

REFERENCES

Alexander, Carol (2008) *Market Risk Analysis: Pricing, Hedging and Trading Financial Instruments*. John Wiley & Sons, Chichester, England.

Andersen, T.G. – Bollerslev, T. (1998) Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts. *International Economic Review*, Vol. 39 (4), 885-905.

Andersen, T.G. – Bondarenko, O. (2007) Construction and Interpretation of Model-Free Implied Volatility. *CREATES Research Paper*, 2007-24.

Arrow, K.J. (1964) The Role of Securities in the Optimal Allocation of Risk-bearing. *The Review of Economic Studies*, Vol. 31 (2), 91-96.

Bakshi, G. – Madan, D. (2006) A Theory of Volatility Spreads. *Management Science*, Vol. 52 (12), 1945-1956.

Bingham, N.H. – Kiesel, R. (2004) *Risk-Neutral Valuation: Pricing and Hedging of Financial Derivatives*. 2nd ed. Springer, London, England.

Birru, J. – Figlewski, S. (2011) Anatomy of Meltdown: The Risk Neutral Density for the S&P 500 in the Fall of 2008. *Journal of Financial Markets*, Vol. 15, 151-180.

Black, F. – Scholes, M. (1973) The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, Vol. 81 (3), 637-654.

Breeden, D.T. – Litzenberger, R.H. (1978) Prices of State-contingent Claims Implicit in Option Prices. *Journal of Business*, Vol. 51 (4), 621-651.

Britten-Jones, M. – Neuberger, A. (2000) Option Prices, Implied Price Processes and Stochastic Volatility. *Journal of Finance*, Vol. 55 (2), 839-866.

Canina, L. – Figlewski, S. (1993) The Informational Content of Implied Volatility. *The Review of Financial Studies*, Vol. 6 (3), 659-681.

Carr, P. – Wu, L. (2009) Variance Risk Premiums. *The Review of Financial Studies*, Vol. 22 (3), 1311-1341.

Chan, W.H. – Cheng, X. – Fung, J.K.W. (2010) Forecasting Volatility: Role of Sampling frequency and forecasting horizon. *Journal of Futures Markets*, Vol. 30 (12), 1167-1191.

Cheng, X. – Fung, J.K.W. (2012) The Information Content of Model-Free Implied Volatility. *Journal of Futures Markets*, Vol. 32 (8), 792-806.

Chernov, M. (2007) On the Role of Risk Premia in Volatility Forecasting. *Journal of Business & Economic Statistics*, Vol. 25 (4), 411-426.

Christensen, B.J. – Hansen, C.S. – Prabhala, N.R. (2001) The Telescoping Overlap Problem in Options Data. *Journal of Financial Economics*, 50, 125-150.

Cox, J.C. – Ross, A.R. (1976) The Valuation of Options For Alternative Stochastic Processes. *Journal of Financial Economics*, Vol. 3, 145-166.

Day, T.E. – Lewis, C.M. (1992) Stock Market Volatility and the Information Content of Stock Index Options. *Journal of Econometrics*, Vol. 52, 267-287.

Diebold, F.X. – Mariano, R. (1995) Comparing Predictive Accuracy. *Journal of Business and Economic Statistics*, Vol. 13, 253-265.

Ederington, L.H. – Guan, W. (2002) Measuring Implied Volatility: Is an Average Better? Which Average? *Journal of Futures Markets*, Vol. 22 (9), 811-837.

Ederington, L.H. – Guan, W. (2005) Forecasting Volatility. *Journal of Futures Markets*, Vol. 25 (5), 465-490.

Elliott, G. – Timmermann, A. (2008) Economic Forecasting. *Journal of Economic Literature*, Vol. 46 (1), 3-56.

Engle, R.F. (1993) Statistical Models for Financial Volatility. *Financial Analysts Journal*, Vol. 49 (1), 72-78.

Figlewski, S. (1997) Forecasting Volatility. *Financial Markets, Institutions & Instruments*, Vol. 6 (1), 1-88.

Figlewski, S. (2008) Estimating the Implied Risk Neutral Density for the U.S. Market Portfolio. In: *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*, eds. T. Bollerslev – J.R. Russell – M. Watson. Oxford University Press, Oxford, England.

Harrison, J.M. – Kreps D.M. (1979) Martingales and Arbitrage in Multiperiod Securities Markets. *Journal of Economic Theory*, Vol. 20, 381-408.

Harrison, J.M. – Pliska S.R. (1981) Martingales and Stochastic Integrals in the Theory of Continuous Trading. *Stochastic Processes and their Applications*, Vol. 11, 215-260.

Jiang, J.J. – Tian, Y.S. (2005) The Model-Free Implied Volatility and Its Information Content. *The Review of Financial Studies*, Vol. 18 (4), 1305-1342.

Lamoureux, C.G. – Lastrapes, W.D. (1993) Forecasting Stock Return Variance: Toward an Understanding on Stochastic Implied Volatilities. *The Review of Financial Studies*, Vol. 6 (2), 293-326.

Liu, X. – Shackleton, M.B. – Taylor, S.J. – Xu, X. (2007) Closed-form transformations from risk-neutral to real-world distributions. *Journal of Banking & Finance*, Vol. 31, 1501-1520.

Lopez, J.A. (2001) Evaluating the Predictive Accuracy of Volatility Models. *Journal of Forecasting*, Vol. 20 (2), 87-109.

Masset, P. (2011) Volatility stylized facts. Working paper, University of Fribourg.

McCracken, M.W. – West, K.D. (2005) Inference About Predictive Ability. In: *Companion to Economic Forecasting*, eds. M.P. Clements – D.F. Hendry. John Wiley & Sons, Chichester, England.

Merton, R.C. (1973) Theory of Rational Option Pricing. *The Bell Journal of Economics and Management Science*, Vol. 4 (1), 141-183.

Newey, W.K. – West, K.D. (1987) A Simple Positive Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, Vol. 55, 703-708.

Phillips, G.M. (2003) *Interpolation and Approximation by Polynomials*. Springer, Secaucus, NJ, USA.

Poon, Ser-Huang (2005) *A practical guide to forecasting financial market volatility.* John Wiley & Sons, Chichester, England.

Rubinstein, M. (1994) Implied Binomial Trees. *Journal of Finance*, Vol. 69 (3), 771-818.

Samuelson, P.A. – Merton, R.C. (1969) A Complete Model of Warrant Pricing that Maximizes Utility. *Industrial Management Review*, Vol. 10, 17-46.

Sundaram, R.K. (1997) Equivalent Martingale Measures and Risk-Neutral Pricing: An Expository Note. *Journal of Derivatives*, Vol. 5 (1), 85-98.

Taylor, S.J. – Yadav, P.K. – Zhang, Y. (2010) The Information Content of Implied Volatilities and Model-Free Volatility Expectations: Evidence on Options Written on Individual Stocks. *Journal of Banking & Finance*, Vol. 34, 871-881.