

**UNIVERSITY OF VAASA**  
**FACULTY OF BUSINESS STUDIES**  
**DEPARTMENT OF ACCOUNTING AND FINANCE**

Samu Kuosmanen  
**INFORMATION CONTENT OF IMPLIED VOLATILITY,  
SKEWNESS AND KURTOSIS**

Master's Thesis in  
Accounting and Finance

VAASA 2006



TABLE OF CONTENTS	PAGE
1. INTRODUCTION	5
1.1. Purpose of the study	7
1.2. Structure of the study	8
2. LITERATURE REVIEW	9
2.1. Implied Volatility	9
2.2. Estimation of Implied Distributions	12
2.3. Information content of Implied Distributions	15
3. OPTION PRICING THEORY	20
3.1. Stock Price Behaviour	20
3.2. Risk-neutral valuation	24
3.3. The Black-Scholes Model	26
3.3.1. Derivation of the model	26
3.3.2. Derivation of the model using risk-neutral valuation	30
3.3.3. Critique against the model	31
3.4. Volatility	32
3.4.1. Volatility smile	32
3.5. Skewness- and Kurtosis-Adjusted Black-Scholes Model	34
4. IMPLIED DISTRIBUTION	39
4.1. Shimko's Method	40
4.2. Corrado and Su Method	40
5. HYPOTHESES	42
6. DATA AND METHODOLOGY	44
6.1. Data	44
6.2. Research process	50
7. EMPIRICAL RESULTS	52
7.1. Results on implied volatility	52
7.2. Results on skewness	56
7.3. Results on kurtosis	58
8. CONCLUSIONS	62
REFERENCES	65



---

**UNIVERSITY OF VAASA****Faculty of Business Studies****Author:** Samu Kuosmanen**Topic of the Thesis:** Information Content of Implied Volatility, Skewness and Kurtosis**Name of the Supervisor:** Petri Sahlström**Degree:** M.Sc. (Economics and Business Administration)**Department:** Department of Accounting and Finance**Major Subject:** Accounting and Finance**Line:** Finance**Year of Entering the University:** 2000**Year of Completing the Thesis:** 2006**Pages:** 68

---

**ABSTRACT**

The purpose of this study is to investigate whether option price implied volatility, skewness and kurtosis are good estimates of realized return distribution. Earlier studies suggest that implied moments, i.e. volatility, skewness and kurtosis, of the distribution do contain some information about future price behavior, but the information is usually biased and exaggerates the importance of past market shocks.

This study employs method introduced by Corrado & Su to obtain estimates of implied volatility, skewness and kurtosis. The data consists of daily close values of DAX index for years 1999-2001. Furthermore, regression analysis is used to compare the information content of implied and history-based estimates to see if implied estimates contain some additional information about future price behavior.

The overall results indicate that implied volatility, skewness and kurtosis do contain some information about the future volatility, skewness and kurtosis, but as the prediction power of these models used in this study is so low, it is difficult to implement this information on predicting the future.

---

**KEYWORDS:** option pricing, Black-Scholes, skewness and kurtosis



## 1. INTRODUCTION

The importance of financial derivatives has increased rapidly in the past 20 years. Most of this success is due to development of accurate pricing models such as Black-Scholes (1973) formula and binomial model by Cox, Ross and Rubinstein (1979). With these models it is easy to calculate option prices if input parameters crucial to option valuation are known. These parameters are stock price, strike price, time to maturity, risk-free interest rate and volatility. All other parameters but volatility can be obtained without much effort because they reflect present time and can be found from electronic- or newspaper quotes. In order to value options, volatility, measuring the uncertainty about the future price behaviour, has to be estimated some way. There is no closed form solution to calculate volatility from e.g. Black-Scholes option pricing model (later B-S model), but instead of that it can be estimated from historical data or it can be solved from market price using iterative methods.

B-S model is based on an assumption that all the options with same underlying asset and same maturity should have equal volatility throughout all the strike prices. Volatility structure can be illustrated by plotting implied volatility as a function of strike price. It usually takes somewhat skewed form, or it can even form a 'smile', meaning that volatility rises always when moving from at-the-money (ATM) to either in-the-money (ITM) or out-of-the-money (OTM). According to Rubinstein's (1994) study the equal volatility assumption did actually hold quite well until the market crash in October 1987. After the market crash the prices of OTM put options, and because of put-call parity also the price of the ITM call options, began to rise as investors feared for further decline in stock prices. It can be observed from the option markets, that the volatility smile has not restored to the flat pre-crash form. In their study Dumas, Fleming and Whaley (1998) found support for the hypothesis that implied volatility patterns exhibit time dependant divergencies.

Because market participants are willing to know markets consensus opinion of future price development, and B-S model gives us tools to estimate markets expectations of the volatility of the underlying asset, implied volatility has been

studied with great interest since the development of B-S model. Shortly after Fischer Black and Myron Scholes (1973) had published their revolutionary option-pricing model, Latané and Rendelman (1976) investigated the relationship between the realised volatility and the implied volatility. They found that the implied volatility significantly outperforms historically estimated volatility in future volatility prediction. Subsequent studies by various researchers have confirmed that implied volatilities are more useful than historical volatility when the realised future volatility is forecasted.

The existence of volatility smile makes the use of implied volatility as a forecast of realised future volatility more complicated. Because volatility changes among the strike price, it is not clear which strike prices should be used in volatility calculation. At the same time when volatility smile complicates the use of implied volatility, it gives us new possibilities to predict the future price behaviour of an asset. The shape of a volatility smile can be used to reveal investors estimates about the probability distribution of the stock price at the maturity. If investors value options with low and high strike prices relatively higher, i.e. they use higher implied volatility for OTM options than ATM options then it implicates that they assume the extreme stock price outcome being more probable than what the lognormal distribution suggests.

Volatility smile can be exploited to obtain risk-neutral densities, which reveal market operators expectations about the stock market returns in the future. Implied probability distribution derived from a volatility smile is a useful tool in examining market participants anticipations of underlying assets future volatility and the direction of probable price movements. If the volatility smile deviates from the flat line, then also the implied probability distribution must deviate from the normal distribution. This non-normality of the implied distribution can be approximated by skewness and kurtosis, third and fourth moments of the distribution.

This study concentrates on the prediction power of implied distributions. It is accomplished by comparing correlations between implied and realised versus historical and realised moments of corresponding distributions. Implied moments are estimated from daily data of German DAX index, covering years

1999-2001. Methods used in the estimation are the very same as used by Corrado and Su (1996) and later by Navatte and Villa (2000).

### 1.1. Purpose of the study

For market participants and policy makers it would be very useful to be able calculate and interpret the market consensus view about predicted direction and volatility of the future asset behaviour. Traders could use implied distributions as a support and additional information for their market decisions and strategies, whereas policy makers could use implied probability distribution functions (PDFs) to forecast how the markets will react on their economical decisions. Two important questions rise when using implied PDFs as a forecast of future market events is discussed: How truthful PDFs estimation methods yield, and, how accurately implied PDFs can predict future realised price behaviour. This study concentrates on the latter issue.

The purpose of this study is to investigate the market participants' ability to predict future price behaviour, or more theoretically expressed does option market implied probability density function predict future underlying asset price behaviour. This is carried out by using implied volatility, skewness and kurtosis as proxy of the distribution. Earlier similar studies have been accomplished using, for example, S&P500 (Corrado & Su 1996) and CAC-400 (Navatte & Villa 2000) data. This study contributes to the current literature using daily data of German DAX index.

If the results indicate that implied distributions have no additional information content and deep-ITM or -OTM options are in that sense overpriced, then the situation can be taken advantage of by selling the overpriced options and neutralising the risks by buying undervalued options. If the results reveal that implied distributions do contain statistically significant information about the future price behaviour, then it can be also very advantageous to use the information in practice. In either case, the results are very interesting.

## 1.2. Structure of the study

This study is constructed so that first the previous research, related to this study, is reviewed in chapter 2. Essential theoretical framework is discussed in chapters 3 and 4, the earlier concentrating mainly on option pricing theory and the latter on implied distributions and the ways to recover them. 5<sup>th</sup> chapter introduces the research hypotheses. Chapter 6 introduces the data and discusses research methods. Empirical results are presented and discussed in Chapter 7. Chapter 8 contains final summary and conclusions of the study.

## **2. LITERATURE REVIEW**

In this chapter the previous research, most important to this study, are reviewed and discussed. The studies under consideration are divided into three categories which describe the focus of these studies. These categories are implied volatility, estimation of implied distributions and the information content of implied distribution. Because these topics are closely related and most of these studies deal with more than one of these topics, it has been difficult to assign some single study to one specific topic. However, these studies have been tried to divide into logical categories on the basis of what issue has had most emphasis on the study and what has the main purpose of the study been.

### **2.1. Implied Volatility**

Information content of implied volatility has been studied by many researchers, for example Beckers (1981), Canina and Figlewski (1993), Jorion (1995) and Fornari & Mele (2001). Most of these studies support the idea that the implied volatility is the best available, yet biased, forecast of the volatility of the future returns. Canina and Figlewski (1993) found opposite results and conclude that implied volatility is a poor predictor of the future realised volatility.

Beckers (1981) investigates the predictive ability of implied standard deviations. The research is done by using regression analysis where actual realised standard deviation acts as a dependent variable and historical and implied standard deviations as explanatory variables. Three differently weighted implied standard deviation measures are used to investigate which options (ITM or ATM) contain most information about the future, but for the current study only the difference between historical and implied volatility matters. The data used in Beckers' study contain CBOE option prices, and because those options are not payout protected, the standard deviations are obtained using procedure that calculates correct value of an unprotected American call option. The results show that in almost every case, the implied standard deviation outperforms the historical standard deviation in predicting the future realised

standard deviation. Furthermore, including both historical and implied standard deviation to the model did increase the  $R^2$ -measure in every sample period. Beckers notes that this increase in  $R^2$  may indicate market inefficiency, because some past information was not revealed in actual option prices.

Canina & Figlewski (1993) were suspicious about widely accepted assumption that implied volatility is better predictor for the future volatility than the historical volatility. They investigated the market's ability to forecast future price volatility by forming simple regression equations for realised volatility where implied volatility and historically measured volatility were used as predictors. The data sample for the study was constructed from daily closing prices of S&P 100 Index options in period 15.03.1983 – 28.03.1987, and it contained total 17 606 observations. Options with fewer than 7 or more than 127 days to expiration and options with more than 10% in- or out-of-the-money were excluded from the data. The results show that the implied volatility is a very poor estimate of the future volatility and it is outperformed even by historically estimated volatility. Canina & Figlewski conclude that they do not believe that small information content of implied volatility originates from irrational option traders, but more probably from the law of supply and demand which might alter the price, or in this case the implied volatility, of some specific type of options. This idea gives rise to the discussion whether the whole strike-volatility structure, or volatility smile, should be taken into account when future volatility is predicted.

In his study, which this study is quite similar to the paper by Canina & Figlewski (1993), Jorion (1995) investigates the information content and the prediction power of currency market implied volatilities. Information content is studied by measuring the implied volatility's ability to forecast subsequent 1-day volatility, while predictive power is tested by regressing the implied volatility on the volatility over the remaining days until the maturity. The actual testing is done by using regression analysis, which was used in many earlier studies on this issue. The data for the study was taken from the Chicago mercantile exchange's closing quotes for currency futures and options on futures. The three most active currencies were chosen, meaning that the data consists of options on Deutsche mark, Japanese yen and Swiss franc. The period

the data covers varies a bit depending on the currency, but it begins around the year 1985 and ends in February 1992. The results show that implied volatilities contain a substantial amount of information about currency movements on the following day, slopes of regression equation for all three currencies being around 0.8, while historical volatilities have slopes as low as 0.3 - 0.4 level. The results for prediction power of implied volatility are alike; Implied volatility outperforms the historically measured volatility significantly. Regression equation slopes on realised volatility are around 0.5 and 0.15 for implied- and historical volatility, respectively. Jorion notes that the results are quite opposite to the Canina & Figlewski study, and suggests that the difference in results may be due to error in estimation of S&P100 implied volatilities.

Fornari & Mele (2001) study how scheduled and unscheduled news affect the implied volatility of options on Italian 10-year bond traded in LIFFE during observation period March 1994 - March 1997. In addition to earlier studies, Fornari & Mele focus on OTM and ITM volatility which they assume to be more sensitive to news releases. For each day three measures for implied volatility are estimated, meaning that only the volatilities for most far-out-of-the-money, most deep-in-the-money and most nearest-at-the-money options are calculated. Finally the research hypothesis, that news resolve uncertainty is tested using regression analysis. The paper provides some evidence that the news may help to resolve uncertainty, i.e. the implied volatility decreases, but most of the change occurring to volatilities of ATM options, while OTM and ITM implied volatilities are affected only by marginally.

Because of volatility smile, it is not obvious which strike price should be used for calculation of implied volatility for these tests. Usually the volatility structure is ignored and only at-the-money implied volatility is used. Obviously this kind of approach can be seen problematic, because it leaves much information out of the scope of study. Furthermore it seems evident that while ATM implied volatility may have some forecast power about future short term volatility it is inadequate in predicting long term volatility. Implied volatility should somehow incorporate the non-flat form of volatility structure to be able to reflect better the market's opinion of future volatility.

In many studies it is suggested that option prices can be seen as the sum of Black-Scholes price plus adjustment terms for skewness and kurtosis. These methods are used in studies from authors like Jarrow and Rudd (1982) who employed generalised Edgeworth expansion, Corrado & Su (1996, 1997) who used Gram-Charlier expansion and Madan & Milne (1994) and Abken et al. (1996) who applied Hermite polynomial expansion. Adding skewness and kurtosis to option pricing model helps us to obtain implied distributions in similar manner as implied volatility is usually obtained. Next chapter discusses more about estimation of implied distributions.

## 2.2. Estimation of Implied Distributions

Among many others, Rubinstein (1994) and Jackwerth and Rubinstein (1996) have studied option prices implied probability distributions and ways to recover them. Earlier, probability distributions of stock market returns have been usually estimated from historical time series, which Jackwerth and Rubinstein (1996:2) comment to be very inappropriate method because, as they say: "it may not capture the probability of extreme events and the events of interest are rare or may not be present in the historical record even though they are clearly possible".

Rubinstein's article Implied Binomial Trees (1994) concentrates on different methods to derive risk-neutral probabilities from simultaneously observed European option prices and using these end-node probabilities to build implied binomial trees. Methods introduced there are Amended Longstaff Method (for original version of Longstaff method, see Longstaff 1990), Shimko's Method (see Shimko 1993) and optimization method. Rubinstein finds Amended Longstaff Method problematic, because the discrete probabilities it yields are very volatile, with values jumping from negative to highly positive, which yields probability distribution very different from lognormal. It is also difficult to get the tail probabilities with this method. In order to acquire tail probabilities options with striking prices from 0 to infinity would be needed, unfortunately this is not the case in the real financial markets and interpolation or extrapolation would be needed to obtain price for options with strike prices

in between of real striking prices. Shimko's method utilises the idea, discovered by Breeden and Litzenberger (1978), that if European options with striking prices from zero to infinity would exist, the maturity date risk-neutral probability distribution could be obtained by calculating the second derivate of price respect to striking price for each option. (Rubinstein 1994)

Jackwerth and Rubinstein (1996) use optimisation procedure introduced in Rubinstein's (1994) article to derive risk-neutral probability distributions from option prices. The data used in this study contains all reported trades and quotes covering S&P 500 European index options and futures traded on CBOE and intraday S&P 500 index levels from April 1968 to December 1993. The data is adjusted for occurring dividends. The stock market crash on October 19, 1987 occurs in data period, and thus Jackwerth and Rubinstein compare implied distributions before and after crash. They find a distinct change in shape between the pre-crash and post-crash distribution. Pre-crash distributions are very close to lognormal distribution, while the post-crash distributions exhibit significant leptokurtosis and left-skewness. Although, there is a significant change in distribution after the crash, it seems that after that skewness and kurtosis stay quite stable at their new post-crash levels. This new form of implied distribution assumes another significant decline in the S&P 500 index far more likely than it did before the crash.

In their study *Skewness and Kurtosis in S&P 500 Index Returns Implied by Option Prices* (1996) Corrado and Su suggest skewness and kurtosis of the return distribution as a source of volatility smile and they derived Skewness- and Kurtosis-Adjusted Black-Scholes option pricing model using Cram-Charlier series expansion of a normal density function. They developed extended Black-Scholes model which incorporates the non-normal skewness and kurtosis in the stock return distribution.

Basically their model yields option value which is equal to Black-Scholes value plus the adjustments for skewness and kurtosis. Like implied volatility, skewness and kurtosis are also parameters which cannot be directly observed from the market and thus must be estimated somehow. Fortunately, this extended Black-Scholes model can be used to simultaneously estimate all three

moments, standard deviation (implied volatility), skewness and kurtosis, from the data.

Main purpose of their study was to investigate, whether extended Black-Scholes model removes pricing biases of deep-ITM and deep-OTM options. The data used in their study consists of CBOE S&P 500 index options, and it is adjusted by subtracting the present value of future dividend payments during the maturity. Using Whaley's (1982) simultaneous equations procedure, Corrado and Su estimate implied volatility for Black-Scholes model and implied volatility, skewness and kurtosis for extended Black-Scholes model for each day. Then they calculate theoretical values for each option using these parameters as input in Black-Scholes and extended Black-Scholes model and furthermore compare these theoretical prices to market prices. The results show that using only Black-Scholes model, on average 89% of theoretical prices are outside the bid-ask spread, while using extended model only 63% of theoretical prices lie outside bid-ask spread. Also the average deviation of theoretical prices from spread narrows from \$0.76 to \$0.40, for Black-Scholes and extended Black-Scholes, respectively. The final conclusion of this study is that the pricing accuracy for deep-ITM and deep-OTM options is significantly improved when terms for skewness and kurtosis are included in the option pricing formula.

In their subsequent study, *Implied Volatility Skews and Stock Return Skewness and Kurtosis Implied by Stock Option Prices* (1997), Corrado and Su use the very same methods to estimate option implied skewness and kurtosis for four actively traded stock options. The research problem is two-folded. First of all, the measures for implied skewness and kurtosis are estimated and explored using the very same methods as in their previous study. Secondly, the pricing accuracy of skewness and kurtosis adjusted option pricing model is compared to the Black-Scholes model. The pricing performance test is accomplished so that the prior day estimates are used as an input to calculate the current day theoretical prices for all options belonging to the same maturity group and then the theoretical prices are compared to the real occurred trades. The data used in their study consists of intraday data of four Chicago Board Options Exchange (later CBOE) actively traded stock option contracts. The data includes stock prices, strike prices and option maturities. Because CBOE stock options are

American style, only call options with no dividends during the maturity are included in study. Empirical results, which are consistent with their previous findings, suggest that the implied distributions tend to be negatively skewed and exhibit positive excess kurtosis relative to a normal distribution. Furthermore, Corrado and Su conclude that the Skewness and Kurtosis Adjusted Black-Scholes model significantly improves the pricing of deep ITM and -OTM stock options.

Because, these earlier studies concerning estimation of implied distributions indicate that the implied return distributions exhibit non-normal skewness and kurtosis, and as Corrado and Su conclude, using implied distribution as an input for option pricing model results more accurate option pricing, it seems that implied distributions are useful and worth estimation. Furthermore, Corrado and Su method for obtaining implied distribution using optimisation procedure seems to be easily adapted to the current study.

Even though implied distributions differ from normal distribution and they can be used in option pricing models, the question about prediction power or information content of implied distribution remains. Next chapter discusses this issue more closely.

### 2.3. Information content of Implied Distributions

Many earlier studies on information content of implied volatility suggest the idea that implied volatility has prediction power about future volatility. This assumption has encouraged researchers to study also the information content of higher moments of the distribution.

Navatte and Villa (2000) investigate if also the higher moments, i.e. skewness and kurtosis, of the distribution contain similar information about the future. They used Skewness- and Kurtosis-Adjusted Black-Scholes option pricing model developed by Corrado and Su (1996) to obtain estimates for implied volatility, -skewness and -kurtosis from European long maturity CAC 40 index options data. Descriptive statistics of the implied moments were consistent with

the empirical results of previous studies; On average, implied skewness was significantly negative and implied kurtosis exceed the kurtosis of normal distribution.

Firstly, they found out that the Skewness- and Kurtosis-Adjusted Black-Scholes option pricing model significantly outperforms the traditional Black-Scholes model in out-of-sample pricing. Secondly, they found that implied moments (volatility, skewness and kurtosis) contain significant amount of information about moments of realised return distribution. Furthermore, they discovered that different shapes of the volatility smiles are consistent with different distribution of the underlying returns.

Weinberg (2001) examines how the risk-neutral implied distributions compare with realised distributions and, in addition to earlier studies, he also tests the information content of implied volatility and skewness against the corresponding moments of daily returns. The daily settlement price data covering years from 1988 to 1999 used in this study is obtained from CME and the contracts examined are on the S&P 500 futures, the Japanese yen/U.S. dollar futures, and the deutsche mark/U.S. dollar futures. First of all, implied distributions are estimated from option prices using volatility smoothing, also known as Shimko's method, and then the goodness of fit of these distributions are compared with the goodness of fit of a lognormal distribution with the same mean and standard deviation. Secondly, the information content of these volatility smile smoothed implied volatility and skewness are tested using regression analysis. The results of distribution comparison show that for foreign exchange series, DM and JY, both the lognormal and implied distribution fit the realised distribution reasonably well, even though the lognormal distribution fits the data better. For S&P 500, the risk-neutral implied distribution does not fit the realised return distribution, but when it is adjusted to risk it fits the realised distribution even better than the lognormal distribution. The results concerning information content of implied volatility, which are consistent with the findings of Jorion (1995), indicate that the volatility smoothing method performs slightly worse in predicting future volatility when compared to ATM Black-Scholes implied volatility. Furthermore, the information content of implied skewness was found out to be poor, actually Weinberg found profit

opportunities available from betting against the market when strong skewness existed. As a possible explanation for poor prediction power, he suggests that because market participants hedge portfolios and therefore they are willing to pay higher price than the true statistical risk for their insurance.

Shiratsuka (2001) examines the information content of implied probability distribution in predicting the future return distribution and price behaviour. First of all, he investigates whether the implied distribution contains information about the future realised stock return distribution. Secondly, he examines whether the implied distribution produces useful information in predicting the future stock price changes. The study employs daily data of Nikkei 225 stock price index options and Japanese Government futures options from mid-1989 to mid-1996. The data contains also the daily estimates for implied volatility, -skewness and -kurtosis.

The forecasting power of implied distribution is tested and compared to the historical return distribution using a simple regression analysis. The results suggest that implied volatility does contain some information about future realised volatility, yet historical volatility is still better predictor. Both implied and historical skewness and kurtosis turned out to contain no useful information about future. These results are consistent with Weinberg (2001). Secondly Shiratsuka uses Granger causality tests to explore whether the shape of implied distribution contains useful information about the future price behaviour. The results indicate that there exists some causality, and the shape of implied distribution provides useful information for predicting future volatility and skewness.

In their paper Risk-Neutral Skewness: Evidence from Stock Options, Patrick Dennis and Stewart Mayhew (2002) study the importance of various factors in explaining the skew in implied distributions. They investigate whether the firm-size, leverage, market risk, trading volume or put/call ratio can explain variations in risk-neutral skewness. Because Dennis & Mayhew are only interested in skewness of the distribution, they use method developed by Bakshi and Madan (2000) which allows to easily compute the risk-neutral skewness and does not rely on any particular option pricing model. The data

used in their study is daily trading data of CBOE covering years 1986 - 1996 and including all strikes and maturities for over 1000 underlying stocks and S&P 500 index. They compute skewness for each day and average these values to obtain weekly skewness measures. Also the explanatory variables, including implied volatility, trading volume, beta, firm leverage and put/call ratio are constructed and calculated from each week. Finally the weekly cross-sectional regressions are calculated. They found out that the market risk, measured as beta, significantly affects the risk-neutral density skewness. Stocks with high beta tend to have more negatively skewed implied distribution especially when the market volatility is high or when the implied distribution for index options is negatively skewed. Also, large market value seem to make the skew negative, while higher trading volume makes the skew more positive. Although, put/call volume ratio, which could intuitively be seen as a market's forecast of future price direction, does not correlate with risk-neutral skew.

In their recent study Anagnou et al. (2002) investigate whether implied probability densities provide unbiased forecast of realised probability densities. They use four different approaches, from which three are parametric and one is non-parametric, to estimate implied distributions for options in S&P 500 and the US Dollar / British Pound from 1986 to 2001. The selected approaches are Generalised Beta Approach, Normal Inverse Gaussian Approach, Two-lognormal mixture and B-Splines. They found out that parametric methods are clearly superior to non-parametric, but none of these methods present an appropriate forecast of the true distribution of the underlying at expiry. Their main conclusion is that the implied distributions provide only biased information about future market dynamics. It also seems that implied volatilities reflect more past shocks than provide accurate forecasts of the future. Therefore, using implied distribution as only forecast would lead to overreaction and wrong conclusions about the future price behaviour.

Most of these studies on information content of implied distribution indicate that implied skewness and kurtosis have a poor prediction power on future skewness and kurtosis. Anagnou et al. (2002) found that implied distributions do contain significant information about future distribution, but are biased and therefore cannot be as only forecast of future return distribution. Shiratsuka

(2001) notes that even though neither implied nor historical higher moments do contain useful information, there exist some causality between the shape of implied distribution and future volatility and skewness.

Poor prediction power has few possible explanations which are closely related to each other; 1. Options are used as insurance and therefore market participants are willing to pay higher price, 2. Market participants have unrealistic expectations, which reflect more past than the future. For example market crash may result in high negative implied skewness, because market participants are expecting the market to continue further down.

Moreover, it seems that implied distributions reflect more past shocks and market participants' fears than realistic predictions of future price behaviour. These fears or possibility of decline are related to market risk, which Dennis and Mayhew (2002) found to significantly affect implied return distribution. Stocks with high beta tend to have more negatively skewed implied distribution than stocks with low beta.

In short, these studies concerning the information content of Implied Distributions indicate that implied distributions or implied moments do contain some information about future price behavior, but the information is usually biased and exaggerates the importance of past market shocks. So, using implied distribution as a future market price predictor may lead to poor trading decisions. It seems that because implied distributions provide exaggerated and overreacted information about market expectations, it may be profitable to bet against the market when implied distribution exhibits high (negative or positive) skewness.

### 3. OPTION PRICING THEORY

#### 3.1. Stock Price Behaviour

It is generally assumed that the stochastic process behind the movement of a non-dividend-paying stock is geometric Brownian motion. The geometric Brownian motion model contains two components, expected return and volatility. (Hull 2000:225) According to the geometric Brownian motion, the change in the stock price during a short time period is

$$(1) \quad \Delta S = \mu S \Delta t + \sigma S \Delta z,$$

where  $S$  = stock price,

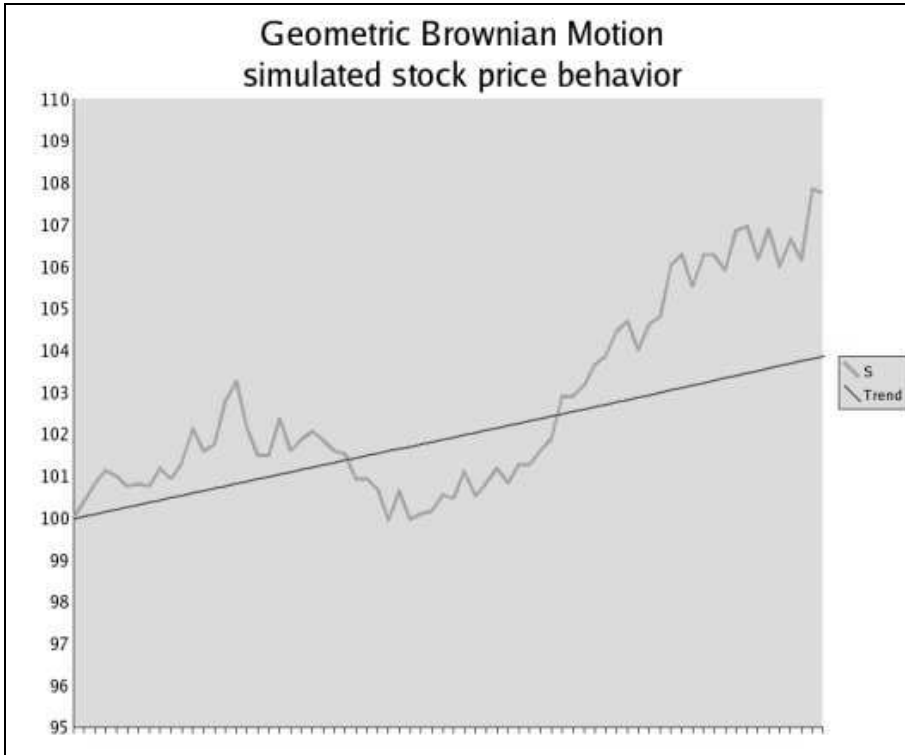
$\mu$  = expected return,

$\sigma$  = volatility,

$\Delta z = \varepsilon \sqrt{\Delta t}$ , where  $\varepsilon$  is a random drawing from normal

distribution,  $\phi(0,1)$

Figure 1 illustrates geometric Brownian motion simulated path followed by a stock. Initial value of stock is 100, expected return used in the simulation is 10% p.a. and stock price volatility is 20%. Hull (2000:227) has presented, that when the volatility equals to zero equation 1 truncates to the form of equation 2. This truncated version of geometric Brownian motion is illustrated as a trendline in figure 1.



**Figure 1.** Geometric Brownian motion simulated stock price behaviour

$$(2) \quad \Delta S = \mu S \Delta t$$

as  $\Delta t \rightarrow 0$

$$(3 \text{ and } 4) \quad dS = \mu S dt \quad \text{or} \quad \frac{dS}{S} = \mu dt$$

and the stock price at the maturity (time T) should equal to

$$(5) \quad S_T = S_0 e^{\mu T}$$

As it can be seen from equation 5, stock price grows at a steady continuously compounded rate  $\mu$ , when the volatility of the stock is zero. Stock price change can be also expressed as a relative to the current stock price. This can be done by dividing equation 1 by the stock price, S.

$$(6) \quad \frac{\Delta S}{S} = \mu \Delta t - \sigma \epsilon \sqrt{\Delta t}$$

In geometric Brownian motion model, both the expected return,  $\mu$ , and the volatility,  $\sigma$ , are assumed constant. From equation 6 we can see that  $\mu\Delta t$  is the expected return over time  $\Delta t$  and, because  $\varepsilon$  is a random drawing from the normal distribution,  $\sigma\varepsilon\sqrt{\Delta t}$  is the only component including stochastic process. This means that the return,  $\Delta S/S$ , is normally distributed with mean  $\mu\Delta t$  and standard deviation  $\sigma\sqrt{\Delta t}$  as expressed in equation 7.

$$(7) \quad \frac{\Delta S}{S} \sim \phi(\mu\Delta t, \sigma\sqrt{\Delta t})$$

Ito's lemma (see Ito 1951) can be used to prove that if the process behind the stock price,  $S$ , is geometric Brownian motion, then the natural logarithm of stock price follows generalised Wiener process and is thus normally distributed. Furthermore, the change in  $\ln S$  during time  $T$

$$(8) \quad d \ln S = \left( \mu - \frac{\sigma^2}{2} \right) dt + \sigma dz .$$

This means that

$$(9) \quad \ln S_T - \ln S_0 \sim \phi \left[ \left( \mu - \frac{\sigma^2}{2} \right) T, \sigma\sqrt{T} \right]$$

and this simplifies further to equation 10, showing that  $\ln S_T$  is normally distributed,

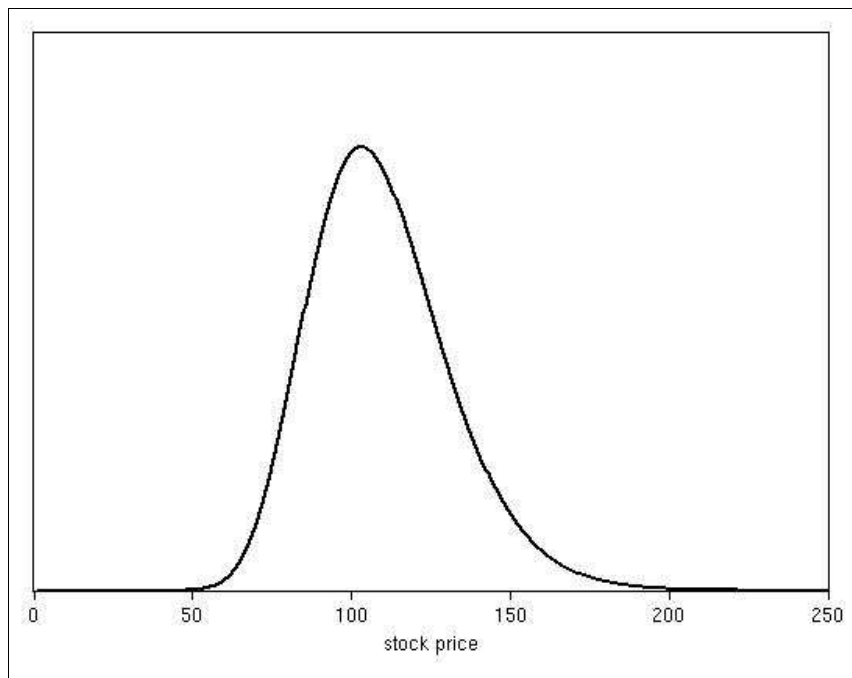
$$(10) \quad \ln S_T \sim \phi \left[ \ln S_0 + \left( \mu - \frac{\sigma^2}{2} \right) T, \sigma\sqrt{T} \right]. \quad \text{Hull (2000:231)}$$

If natural logarithm of a variable is normally distributed, then the variable itself is lognormally distributed. Because equation 10 shows that  $\ln S_T$  is normally distributed, then the stock price at time  $T$  must be lognormally distributed. From the properties of lognormal distribution we know that it can take any value between zero and infinity (see, e.g., Cox and Rubinstein 1985:201-204 and

Smith and Merceret 2000). Furthermore, from equation 10 and the properties of lognormal distribution, we know that the expected value of  $S_T$ ,  $E(S_T)$ , is

$$(11) \quad E(S_T) = S_0 e^{rt}$$

Terminal date stock price distribution implies the probabilities of stock price ending at certain level at the maturity. Figure 2 illustrates the probability distribution of stock price at terminal date. Distribution in figure 2 is calculated with initial stock price of 100, volatility of 20%, interest rate 5%, time to maturity 1 year.



**Figure 2.** Probability distribution of terminal date stock price.

Geometric Brownian motion is only one suggestion for stock price behaviour. It is problematic because of assumptions behind the model. It assumes that the volatility of returns is constant throughout the maturity of an option and that returns are normally distributed, which both are evidently untrue. This gives a rise to a phenomenon called volatility smile.

Other stock price behaviour models are models involving jumps and models with stochastic volatility. Jump model is used in option pricing models such as

The Pure Jump Model by Cox, Ross and Rubinstein (1979) and The Jump Diffusion Model by Merton (1976). Stochastic volatility model is discussed for example in a study by Hull and White (1987).

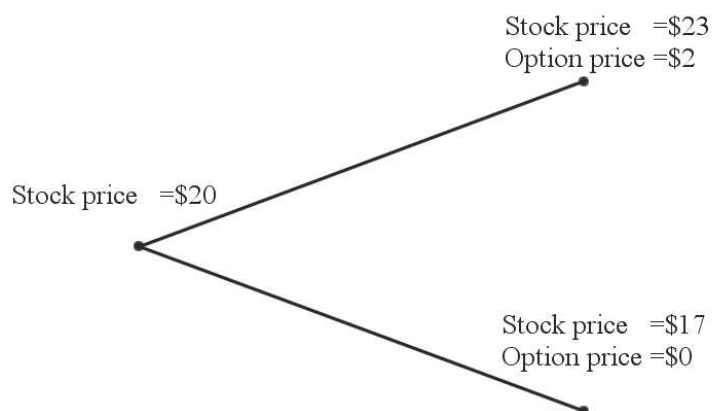
### 3.2. Risk-neutral valuation

In a risk-neutral world all individuals are risk-neutral, and therefore do not require compensation for risk beared. Furthermore, the expected return on all investments equals to the risk-free interest rate and the value of an asset is its expected future price discounted at the risk-free rate. The expected stock price,  $E(S_T)$ , at time T can be expressed as following, (Hull 2000:205)

$$(12) \quad E(S_T) = S_0 e^{rT}.$$

Consider that we have a situation where a stock is currently trading at \$20 and after a one period (say 1 month) the price of the stock can be either \$23 or \$17. In Figure 3, the situation is illustrated with a one-step binomial tree, which is a simplified model of stock price behaviour. According to the equation 12, the expected return on the stock should be equal to risk-free interest rate (say 5% continuous compounding), so the expected stock price after 1 month,  $E(S_1)$ , is

$$(13) \quad E(S_1) = \$20 e^{0.05 \times \frac{1}{12}} \approx \$20.08.$$



**Figure 3.** One Step Binomial tree

As earlier mentioned, after 1 month, the stock price can be either \$23 or \$17 and now we know that the expected value of the stock after 1 month is \$20.08. Now we can calculate the risk-neutral probabilities of up- and down-movements; define  $p$  as the risk-neutral probability of an up-movement. Expected price of a stock is, simply put, the probability weighted average of the possible outcomes and in risk-neutral world it must equal to equation 12. (Hull 2000:206) Example calculation of the probabilities is presented below

$$(14) \quad 23p + 17(1 - p) = 20e^{0.05 \times \frac{1}{12}}$$

$$(15) \quad 6p = 20e^{0.05 \times \frac{1}{12}} - 17$$

$$(16) \quad p = 0.5139$$

Now we can use these risk-neutral probabilities to value an option on the stock. Consider a situation, where we have a European call option with strike price \$21 and 1 month maturity. The option values at the final node are presented in figure 3. At the end of one month, the call option has a 0.5139 (this is the very same as the probability of a stock moving up) probability of being worth \$2 and a 0.4861 (=1-0.5139) probability of being worth zero. The expected value of the option therefore is,

$$(17) \quad 0.5139 \times 2 + 0.4861 \times 0 = 1.0278 .$$

To obtain the risk-neutral value of the option today the expected value should be discounted using the risk-free interest rate. The option is today worth of

$$(18) \quad 1.0278e^{-0.005 \times \frac{1}{12}} = 1.0235 .$$

Using binomial tree, we obtain probabilities of stock price ending at large variety of levels. Of course, one-step binomial tree is extremely simplified model of the stock price behaviour and using it we obtain probabilities for only two outcomes, which are obviously inadequate. But when we add more and more steps, the probability density function becomes more and more accurate and resembles more and more the lognormal density function. This density is

the risk-neutral probability density, and it indicates the probability of the stock price ending at some certain level.

Actually, in order to value an option, we do not have to construct the whole binomial tree. We need only the appropriate risk-neutral probability density, about which is discussed more detailed later in chapter 3.3.2.

### 3.3. The Black-Scholes Model

The Black-Scholes Model, derived by Black and Scholes (1973) expanded by Merton (1973), has played a major role in the development of modern financial derivative markets. Rubinstein (1994) praises the model being one of the most successful models in the social sciences and perhaps even the most widely used formula throughout the human history. It is also generally accepted as a benchmark model, which alternative models are tested against (Nikkinen 2001:9). The importance of this model was recognised all over in the academic world in 1997, when the Nobel prize in economics was awarded to Myron Scholes and Robert Merton (see Hull 2000).

#### 3.3.1. Derivation of the model

The Black-Scholes option-pricing model is derived using the Black-Scholes-Merton differential equation, which is an equation that every derivative on non-dividend paying stock must satisfy in order to fulfil the no-arbitrage assumptions. It is based on an assumption of ideal market conditions and a possibility to create a riskless position of stock and options. Riskless portfolio can be created because both, the stock and the option price, are dependent on the same source of uncertainty, the stock price movement. In a short period of time, price of a call (put) option is perfectly positively (negatively) correlated with the stock price. So, it is possible to create a riskless portfolio of options and a stock, because loss or gain in stock is always offset by an equal gain or loss in option position. Option sensitivity to the stock price is not constant over time and therefore this portfolio is riskless only for a very short period of time and it must be constantly rebalanced to remain riskless for longer period, but during

that short time it must yield risk-free interest rate in order to satisfy the no-arbitrage assumptions. (Hull 2000:224)

Hull(2000:245) has listed assumptions behind the Black-Scholes-Merton differential equation as following:

1. The stock price follows the geometric Brownian motion.
2. The short selling of securities with full use of proceeds is permitted.
3. There are no transaction costs or taxes. All securities are perfectly divisible.
4. There are no dividends during the life of the derivative.
5. There are no arbitrage opportunities.
6. Security trading is continuous.
7. The risk-free rate of interest is constant and the same for all maturities.

To understand the derivation of Black-Scholes-Merton differential equation we have to concentrate on the behaviour of the underlying stock price. In this model it is assumed that the stock price follows geometric Brownian motion, as shown in equation 1 and again in equation 19 below

$$(19) \quad \Delta S = \mu S \Delta t + \sigma S \Delta z .$$

It is showed (see Hull 2000:246 for example) that if  $f$  is a price of a derivative on  $S$ , then using Ito's lemma we find out that process followed by derivative is

$$(20) \quad \Delta f = \left( \frac{\partial f}{\partial S} \mu S + \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \right) \Delta t + \frac{\partial f}{\partial S} \sigma S \Delta z ,$$

where  $\Delta S$  and  $\Delta f$  are the changes of stock price and derivative price over very short time  $\Delta t$ , respectively.

Wiener process behind  $f$  and  $S$  is the same; this means that  $\Delta z$  in both equations is the same. This fact can be used to eliminate the uncertainty by creating a portfolio containing

-1 derivative

and  $+\frac{\partial f}{\partial S}$  shares.

When this kind of portfolio is created, then the value of the portfolio,  $\Pi$ , is

$$(21) \quad \Pi = -f + \frac{\partial f}{\partial S} S$$

and the change in the value of the portfolio,  $\Delta\Pi$ , during a small interval of time,  $\Delta t$ , is

$$(22) \quad \Delta\Pi = -\Delta f + \frac{\partial f}{\partial S} \Delta S .$$

When  $\Delta f$  and  $\Delta S$  in equation 22 are replaced with equations 19 and 20 then the equation can be expressed as follows

$$(23) \quad \Delta\Pi = \left( \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \right) \Delta t .$$

As we can see, equation 23 does not contain stochastic process  $\Delta z$ . Because of this the portfolio must be riskless during time  $\Delta t$  and as the portfolio is riskless, it must yield exactly the same risk-free return as other risk-free securities do. This is presented in equation 24. If it were not true, arbitrageurs could make risk-free profits by selling security/portfolio with smaller return and using the proceeds to buy higher yielding security/portfolio. (Hull 2000:247)

$$(24) \quad \Delta\Pi = r\Pi\Delta t ,$$

where  $r$  is the risk-free interest rate. Substituting  $\Delta\Pi$  and  $\Pi$  in equation 24 with equations 23 and 21 yields

$$(25) \quad \left( \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \right) \Delta t = r \left( f - \frac{\partial f}{\partial S} S \right) \Delta t ,$$

which simplifies to the Black-Scholes-Merton differential equation presented below.

$$(26) \quad \frac{\partial f}{\partial t} + rS \frac{\partial f}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} = rf$$

Solution for Black-Scholes-Merton differential equation can be found for all kind of different derivatives with  $S$  as an underlying variable. Solution of the equation depends of the boundary conditions, which define the range derivative price must lie within. Taken a European call option for example, key boundary condition at maturity of an option is

$$(27) \quad f = \max(S - X, 0).$$

Famous Black-Scholes option pricing model was developed, when Black and Scholes (1973) found solution for European options for Black-Scholes-Merton differential equation. They found out that for boundary conditions of European call option the differential equation simplifies to similar form as the heat-transfer equation in physics and therefore they could use its solution by Churchill (1963) to solve the differential equation. For European put and call options, the solutions which satisfy the Black-Scholes-Merton differential equation are presented below. These are the Black-Scholes option pricing formulas for put and call options

$$(28) \quad c = S_0 N(d_1) - Xe^{rT} N(d_2)$$

$$(29) \quad p = Xe^{-rT} N(-d_2) - S_0 N(-d_1)$$

where

$$d_1 = \frac{\ln\left(\frac{S_0}{X}\right) + \left(r + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}$$

and

$$d_2 = d_1 - \sigma\sqrt{T}.$$

### 3.3.2. Derivation of the model using risk-neutral valuation

Cox and Ross (1976) were first to prove that the Black-Scholes formula can be also derived using the risk-neutral valuation. In proof it is supposed that the stock price follows the geometric Brownian motion, from which follows that the stock price, and because of that, also the risk neutral probability density function, is lognormally distributed.

If  $g(S)$  is defined as a probability density function of  $S$ , then it follows that the expected value of a European call option at the maturity equals to the integral of how much the stock price exceeds the strike price times the probability. This relation is presented in the following

$$(30) \quad \hat{E}[\max(S_T - X, 0)] = \int_X^{\infty} (S - X) g(S) dS .$$

By solving the integral presented in right-hand side of the equation 30 Cox and Ross(1976) obtained their key result, which proved that if the stock price,  $S$ , is lognormally distributed and the standard deviation of  $\ln S$  is  $\sigma$  then

$$(31) \quad E[\max(S - X, 0)] = E(S)N(d_1) - XN(d_2) ,$$

where  $N(x)$  is the cumulative probability distribution and  $d_1$  and  $d_2$  are the very same as presented in equation 29.

As we know, the present value of an option is its expected value discounted at the risk-free interest rate. Value of a European call option,  $c$ , can therefore be expressed as

$$(32) \quad c = e^{-rT} \hat{E}[\max(S_T - X, 0)] .$$

In risk-neutral world the expected value of a stock at the maturity,  $E(S_T)$ , equals to  $S_0 e^{rT}$  (see equation 12). When the key result of Cox and Ross (1976) is applied to equation 32, the value of the European call option,  $c$ , can be expressed as

$$(33) \quad c = e^{-rT} [S_0 e^{rT} N(d_1) - XN(d_2)]$$

or

$$(34) \quad c = S_0 N(d_1) - X e^{-rT} N(d_2)$$

and as can be seen, equation 34 is the very same as the BS model for call options, already presented in equation 28. The terms in equation 33 can be interpreted as following:  $N(d_2)$  represents the probability that the option will be exercised in risk-neutral world and  $S_0 N(d_1) e^{rT}$  is the expected value of the stock if it exceeds the strike price and is zero otherwise. (Hull:251)

It has to be noted that BS model and equation 34 are assuming that the stock price is lognormally distributed. If this does not hold, then the option prices given by the BS model are biased and differ from the market prices. If the very same model is used to obtain values for implied volatility, then this pricing error can be observed as a volatility smile, which is discussed more detailed in chapter 3.4.1.

If the probability density function of the stock price is non-normal, then we just have to use the proper probability density function as an input in equation 30. Actually the probability density function of the stock price,  $g(S)$ , is the most important parameter affecting the value of an option in the risk-neutral valuation model.

### 3.3.3. Critique against the model

According to the empirical studies (see Bates 1996), theoretical prices for OTM and ITM options given by the Black-Scholes model often seriously differ from the prices observed in the market. Especially the deep OTM and -ITM option prices are heavily biased.

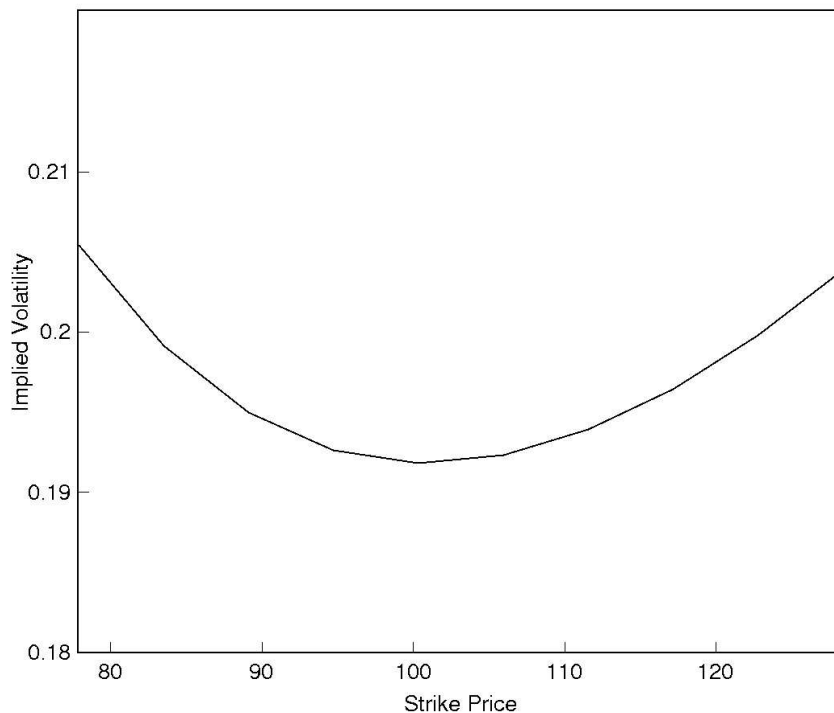
### 3.4. Volatility

Volatility,  $\sigma$ , is one of the crucial parameters needed by the Black-Scholes model to value options. It measures the uncertainty concerning the returns provided by the stock. It is also the only parameter which can not be easily observed and must therefore be estimated, for example from historical data or directly from market prices. If this market implied volatility is needed, then some iterative methods must be used. Stock price volatility is usually somewhere between 20% and 40%, depending on a riskiness of a company. (Hull 2000:241)

Jackwerth and Rubinstein (1995) say that “Historically measured volatility varies significantly over different time intervals; and second this can be a poor predictor of subsequent implied volatility.”

#### 3.4.1. Volatility smile

The Black-Scholes model assumes the volatility to be constant throughout all the striking prices. This is consistent with the idea that stock price follows the geometric Brownian motion and changes are lognormally distributed. Evidently neither of these assumptions is true; Hull (1993) and Nattenburg (1994) found out that the stock returns exhibit non-normal skewness and kurtosis. These non-normalities of the distribution cause volatility smile, meaning that implied volatilities of deep-out-of-the-money and deep-in-the-money options differ from those of at-the-money options. Volatility smile can be illustrated by calculating implied volatilities for different striking prices and then plotting them as a graph. An example of a volatility smile is presented in figure 4 below.

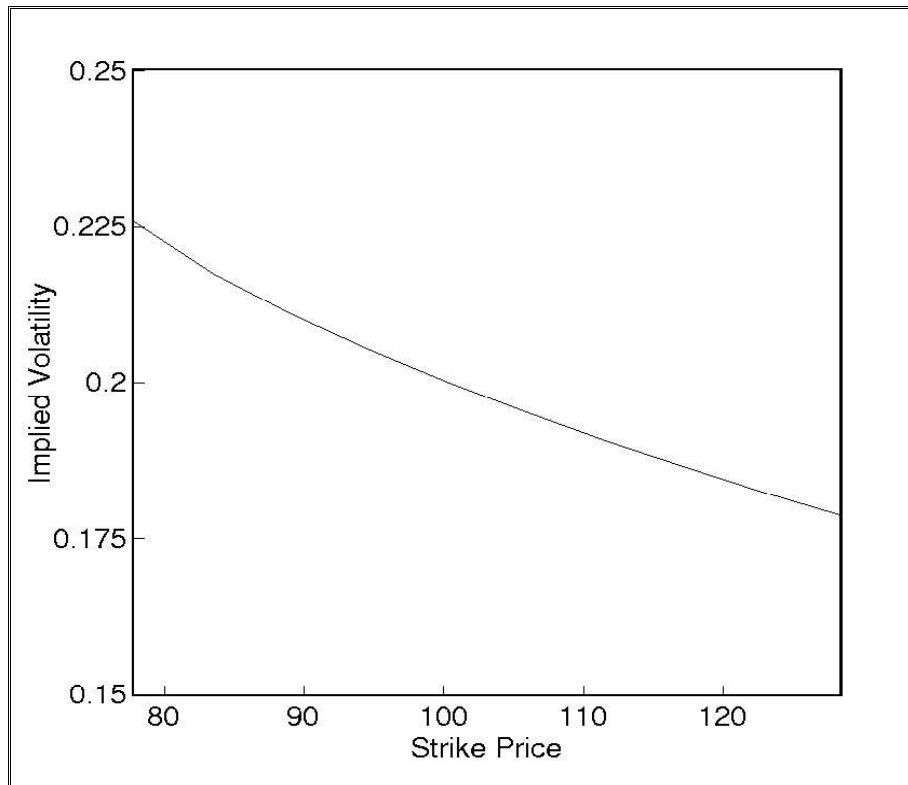


**Figure 4.** An example of volatility smile

The term volatility smile usually refers to a special form of volatility smile, where implied volatility is lowest for ATM options rising towards both, low and high, striking prices. Because options with extreme strike prices are valued relatively higher by using higher volatility for those options, than ATM options, this kind of smile implies that market operators expect extreme movements to be more probable than what would be predicted on the basis of lognormal distribution. This implied distribution exhibits more kurtosis, i.e. fatter tails and more peaked, than the lognormal distribution, meaning that both small and large price changes are more probable and medium changes are less likely than with the lognormal distribution. (Hull 2000:437)

Equity options usually exhibit another special case of volatility smile, called volatility skew. It means that the volatility is relatively high for options with low striking price decreasing gradiently as the striking price increases; this is presented in figure 5. The implied distribution consistent with volatility skew is somewhat negatively skewed and has higher kurtosis than for the lognormal distribution. Rubinstein (1994:775) gives “crash-o-phobia” as one possible explanation for this phenomenon. The idea is that traders fear that a crash

similar to October 1987 could happen again and therefore they value deep OTM put (and because of put-call-parity also the ITM call) options very high relative to ATM options. It seems quite valid explanation, because the Black-Scholes model did hold satisfactory well until the mid 1980's (see Rubinstein 1994).



**Figure 5.** An example of volatility skew

### 3.5. Skewness- and Kurtosis-Adjusted Black-Scholes Model

Corrado and Su (1996) suggest skewness and kurtosis of stock return distribution as a primary source of volatility smile and develop an extended Black-Scholes model, which incorporates with these non-normalities of the return distribution. Their formula consists of the Black-Scholes option pricing formula and the adjustment terms for non-normal skewness and kurtosis. Adjustment terms for skewness and kurtosis are based on a Gram-Charlier series expansion of the standardised normal density function. Corrado and Su truncated the Gram-Charlier series expansion so that only the first four moments are included and that truncated risk-neutral density function,  $g^*(z)$ , can be expressed in the following form:

$$(35) \quad g^*(z) = n(z) \left[ 1 + \frac{\mu_3}{3!} (z^3 - 3z) + \frac{\mu_4 - 3}{4!} (z^4 - 6z^2 + 3) \right],$$

$$\text{where} \quad z = \frac{\ln(S_t / S_0) - (r - \sigma^2 / 2)T}{\sigma\sqrt{T}}$$

and  $n(z)$ ,  $\mu_3$  and  $\mu_4$  are the standard normal density function, skewness and kurtosis respectively. Equation 35 simplifies to the standard normal density function as skewness,  $\mu_3$ , equals to 0 and kurtosis,  $\mu_4$ , equals to 3. Using truncated density function Corrado and Su (1996) derived the approximate formula, presented in equation 36, which can be used to value European call options. The skewness- and kurtosis adjusted option pricing formula for European call options is following:

$$(36) \quad C_{GC} = C_{BS} + \mu_3 Q_3 + (\mu_4 - 3) Q_4,$$

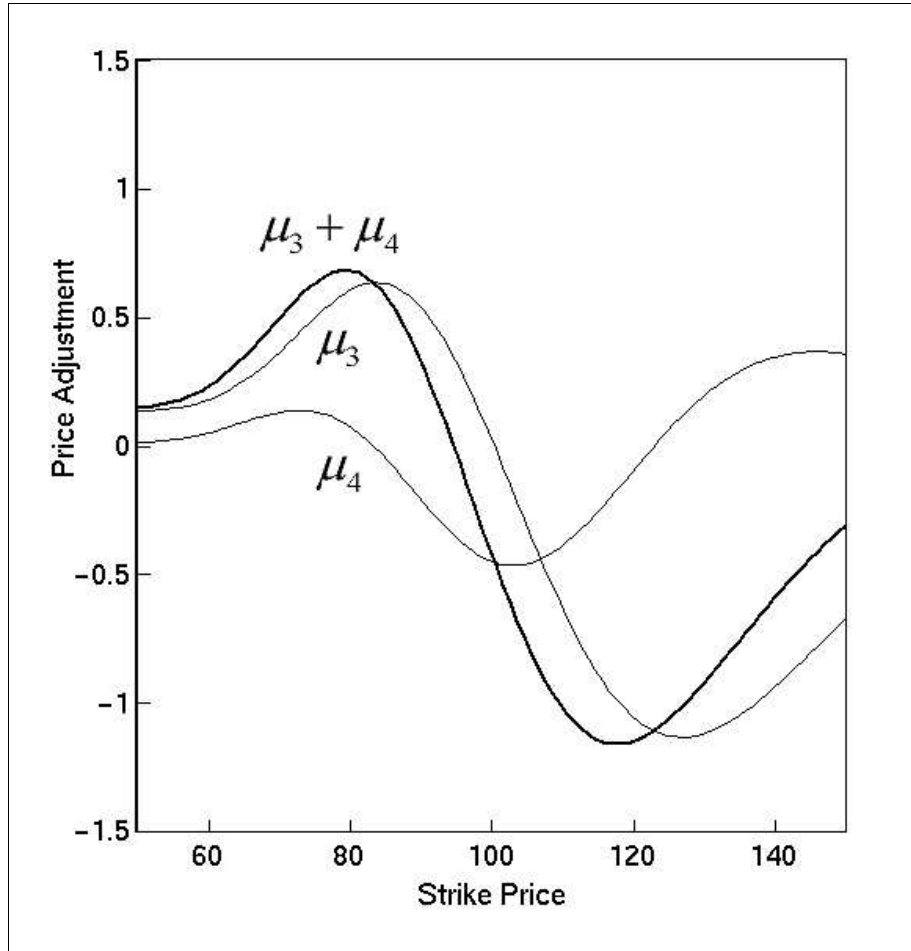
where

$$Q_3 = \frac{1}{3!} S_t \sigma \sqrt{T} \left[ (2\sigma\sqrt{T} - d) n(d) - \sigma^2 T N(d) \right],$$

$$Q_4 = \frac{1}{4!} S_t \sigma \sqrt{T} \left[ (d^2 - 1 - 3\sigma\sqrt{T}(d - \sigma\sqrt{T})) n(d) - \sigma^3 T^{3/2} N(d) \right],$$

$$d = \frac{\ln(S_0 / K) + (r + \sigma^2 / 2)T}{\sigma\sqrt{T}}.$$

The terms  $Q_3$  and  $Q_4$  measure the effects of non-normal skewness and kurtosis on the option price and the effects are also illustrated in Figure 5 below. It should be noted, that when the risk-neutral distribution is lognormal (i.e. skewness=0 and kurtosis=3) the Skewness- and Kurtosis-adjusted Black-Scholes model yields to the very same result as the original Black-Scholes model ( $C_{GC} = C_{BS}$ ). Figure 6 illustrates the price adjustments due to non-normalities. (Corrado and Su 1996:179)



**Figure 6.**  $S = 100$ ,  $\sigma = 0.2$ ,  $r = 0.05$ ,  $T = 1$  year,  $\mu_3 = -0.5$ ,  $\mu_4 = 3.5$ .

Equation 36 can be used to value European put options as well. This can be done by applying put-call-parity relationship between put and call options which must hold in order to satisfy no-arbitrage assumptions. Using put-call-parity a skewness and kurtosis corrected price of a put option can be denoted as follows: (Navatte and Villa 2000:44)

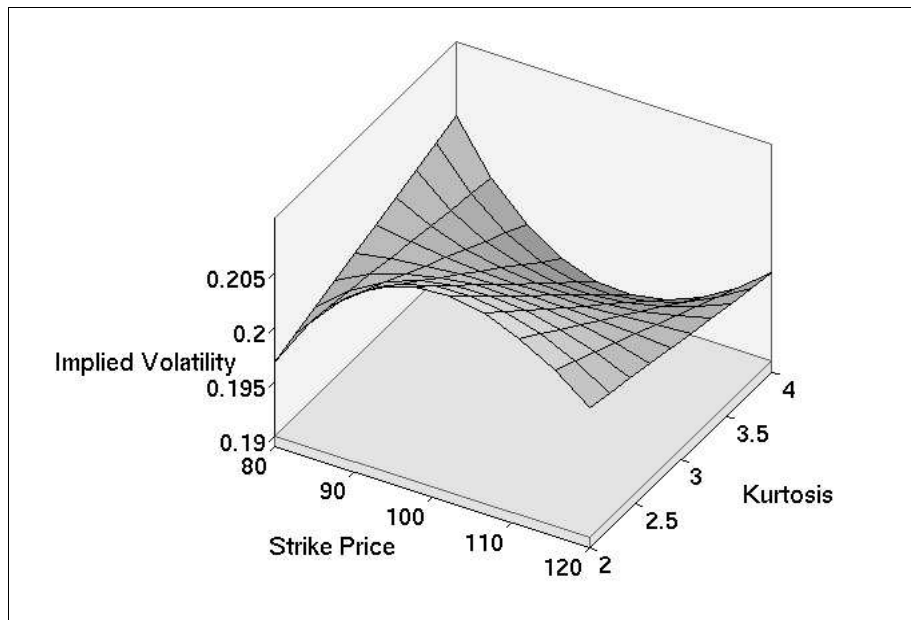
$$(37) \quad p_{GC} = c_{GC} - S_t + Ke^{-rT}.$$

Even though, neither the Black-Scholes model or Skewness and Kurtosis Adjusted Black-Scholes model can be inverted so that the implied volatility could be directly calculated, some approximation functions for implied volatility can be derived. Navatte and Villa (2000) derived an approximation function for implied volatility with following input parameters: volatility,  $\sigma$ ,

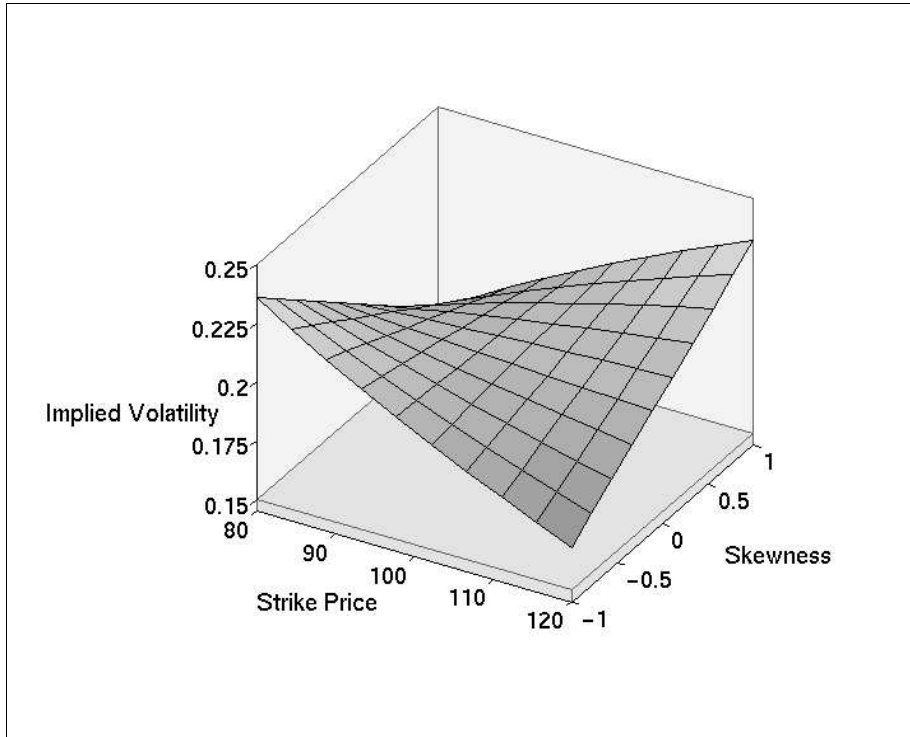
option moneyness,  $x$ , time to maturity,  $T$ , skewness,  $\mu_3$  and kurtosis,  $\mu_4$ . The approximation function for implied volatility,  $\sigma_I$ , is following:

$$(38) \quad \sigma_I = \sigma + \frac{\mu_3}{3!} \sigma \left[ (2\sigma\sqrt{T} - d) - \sigma^2 T \frac{N(d)}{n(d)} \right] + \frac{\mu_4}{4!} \sigma \left[ (d^2 - 1 - 3\sigma\sqrt{T}(d - \sigma\sqrt{T})) - \sigma^3 T^{3/2} \frac{N(d)}{n(d)} \right]$$

Equation 38 can be used to visually illustrate how much skewness and kurtosis affect implied volatility and how does the volatility smile change when skewness or kurtosis changes. Using this approximation, Navatte and Villa (2000) found out that different shapes of volatility smiles are consistent with different return distributions. For example positive excessive kurtosis, which means distribution with “fat tails”, is consistent with symmetric volatility smile (see Figure 3). Effects of kurtosis are more closer investigated in figure 7. When kurtosis alone leads to symmetric volatility smile, skewness creates skewed volatility smiles, or volatility skews. Positive (negative) skewness of the return distribution is consistent with volatility skew, meaning that the implied volatility rises as the moneyness increases (decreases), as illustrated in figure 8.



**Figure 7.**  $S = 100$ ,  $\sigma = 0.2$ ,  $r = 0.05$ ,  $T = 1$  year and  $\mu_3 = 0$ .



**Figure 8.**  $S = 100$ ,  $\sigma = 0.2$ ,  $r = 0.05$ ,  $T = 1$  year and  $\mu_4 = 3$ .

#### 4. IMPLIED DISTRIBUTION

As earlier discussed, when the risk-neutral valuation is used to value options the most important thing to know is the appropriate probability density function. There are many ways to obtain this density function. As the estimate for volatility, similarly the probability distributions of stock market returns can be estimated from historical time series. The problem with this method is that the probability distribution based on past price behaviour reflects the past and not the future, also it may be that the extreme events are not present in the historical record even though it is obvious that they are possible. Jackwerth and Rubinstein(1996) take a good example about the possibility of extreme events: "...the stock market crash of October 1987. Following the standard paradigm, assume that stock market returns are lognormally distributed with an annualised volatility of 20% (near its historical realisation). On October 19, 1987, the two month S&P 500 futures fell 29 percent. Under the lognormal hypothesis, this is a -27 standard deviation event with probability  $10^{-160}$ , which is virtually impossible. Nor is October 1987 a unique refutation of lognormal hypothesis. Two years later, on October 13, 1989, the S&P 500 index fell about 6 percent, a -5 standard deviation event. Under the maintained hypothesis, this has a probability of 0.00000027 and should occur only once in 14756 years." (Jackwerth and Rubinstein (1996):1611-1612)

Alternatively the probability distribution function can be directly implied from the option markets. This means that option data can be used to obtain this so-called risk-neutral distribution or implied distribution. It is obvious, that the implied distribution is not necessarily the same as the realised distribution. Implied distribution is useful because it can be regarded as the market participants consensus forecast of the future price of the underlying, at least it is forward looking estimate in contrast to the probability distribution estimated from historical prices. (see Anagnou, Bedendo, Hodges and Tompkins 2002:2)

Next chapter introduces some methods to derive implied distributions.

#### 4.1. Shimko's Method

In their early study Breeden and Litzenberger (1978) presented the idea, that the risk-neutral probability distribution function can be obtained by calculating the second derivative of each option with respect to its striking price. In order to solve the entire distribution, the data should contain values of european options with the same maturity, the same underlying asset and the strike prices from 0 to infinity.

Unfortunately, in reality we observe market values only for a quite narrow range, which definitely is not even close to from zero to infinity range. Shimko (1993) presented a way to use Breeden's and Litzenberger's idea, which does not need dense strike price continuum. Shimko suggested fitting a smooth curve to the volatility smile plot, in order to obtain interpolated volatility values for every strike price. Then, using the Black-Scholes formula, the option price could be presented as a continuous function of strike price. Finally, taking the second derivative of the option price function, the implied risk-neutral distribution between the lowest and the highest strike price is a result.

Rubinstein (1994) comments Shimko's method being appropriate in a way that it results lognormal risk-neutral probability distribution when the volatility smile used is flat.

#### 4.2. Corrado and Su Method

In their study 'Skewness and kurtosis in S&P 500 index returns implied by option prices' Corrado and Su (1996) derive Skewness- and Kurtosis-adjusted Black-Scholes model (see chapter 2.5 for the derivation) and use maximum likelihood method to simultaneous estimate implied-volatility, -skewness and -kurtosis via the model. The estimation method of Corrado and Su does not assume any specific distribution behind the risk-neutral density function, it just estimates the standard deviation, skewness and kurtosis of the distribution.

The parameters are obtained by minimising the following sum of squares with respect to  $ISD$ ,  $ISK$  and  $IKT$ :

$$(39) \quad \min_{ISD, ISK, IKT} \sum_{j=1}^N [C_{OBS,j} - (C_{BS,j}(ISD) + ISK \times Q_3 + IKT \times Q_4)]^2$$

The output values for  $ISD$ ,  $ISK$  and  $IKT$  represent estimates for implied volatility, implied skewness and implied kurtosis, respectively. (Corrado and Su 1996:184.)

## 5. HYPOTHESES

The present study concentrates on the information content of implied distribution in respect to the future price distribution. In this study, the information content actually means the prediction power and thus it is investigated if the implied distribution predicts the shape of future distribution.

Actual distributions are not compared to each other, but implied moments (volatility, skewness and kurtosis) are used as proxy of the implied distribution. So, instead of comparing implied and realized distributions we compare three moments of those distributions; implied- vs realized volatility, implied- vs realized skewness and implied- vs realized kurtosis.

Comparing only implied and realized moments does not make any sense, because without any reference value it is impossible to say whether some implied moment contains any relevant information about future realized moment. Therefore we use historical moments as reference value and investigate if they perform better in predicting future realized moments than the implied moments. In order to be able to say that implied moments contain additional information about the future, the prediction power of implied moments should be better than the prediction power of historically measured moments.

Because implied moments (volatility, skewness and kurtosis) are used as proxy of actual distribution and thus we have three different measures to be tested, we have to form three different research hypotheses, one for every moment. The research hypotheses to be tested can be expressed as following:

- H<sub>1</sub>        Implied volatility contains additional information about the future volatility, i.e. the prediction power of implied volatility in respect to the future realised volatility is better than the prediction power of historical volatility.
  
- H<sub>2</sub>        Implied skewness contains additional information about the future skewness, i.e. the prediction power of implied skewness in respect to

the future realised skewness is better than the prediction power of historical skewness.

H<sub>3</sub> Implied kurtosis contains additional information about the future kurtosis, i.e. the prediction power of implied kurtosis in respect to the future realised kurtosis is better than the prediction power of historical kurtosis.

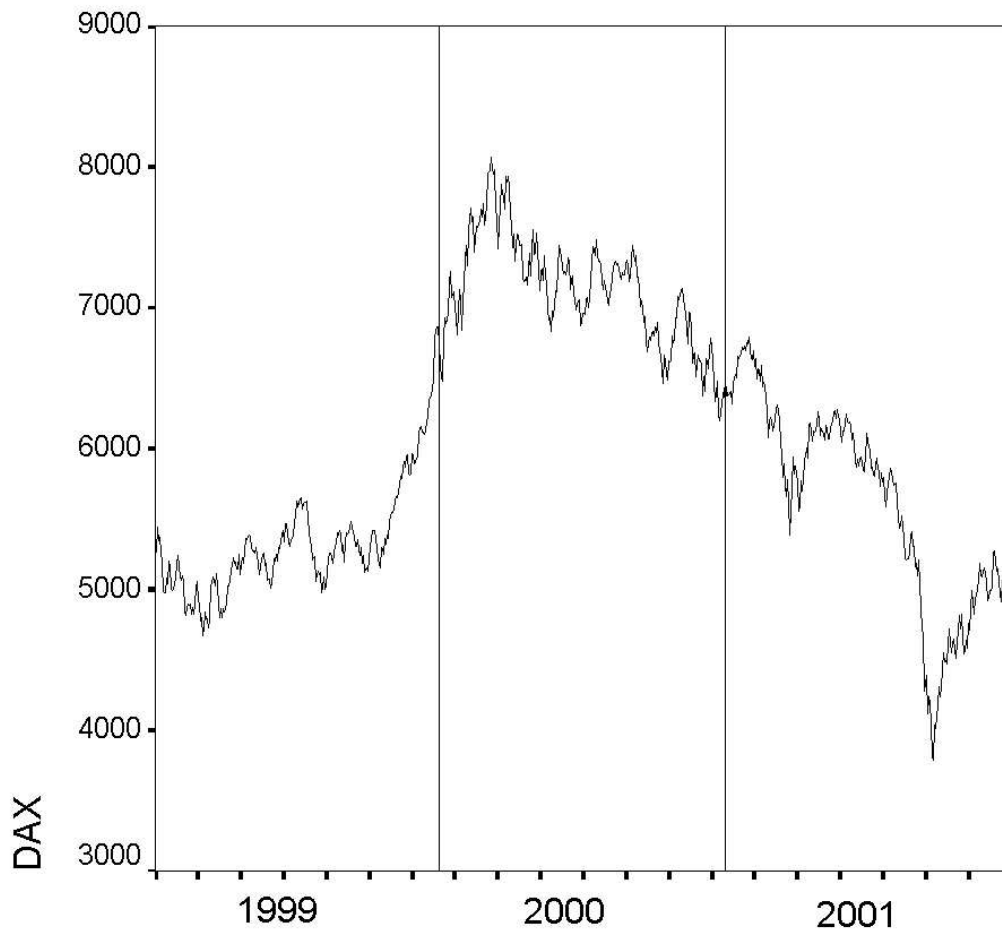
## 6. DATA AND METHODOLOGY

### 6.1. Data

The data in this study comprises of three years daily data of German stock index DAX and is provided by Estlander & Rönnlund Financial Products Ltd. The time period of the data is 4.1.1999 – 28.12.2001. Each day, the data contains end-of-the day prices for DAX index and options on the index.

DAX index is so called performance index, so it has already been adjusted for capital changes and dividends. The daily ending prices of DAX index are so called daily settlement prices meaning that it is the price of the last trade that occurred during the last 15 minutes of trading on an exchange trading day. If it is not possible to determine a price under these conditions or if the price so determined does not reflect the true market conditions, Deutsche Börse will set the settlement price. (Deutsche Börse 2002)

Figure 9 illustrates the behaviour of DAX index during the observation period. As it can be seen from the graph the market conditions have been very volatile during the period. Most of the year 1999 the index stayed at steady 5000 level until the end of the year when it rose steeply to 8000 level in march 2000. After peaking the all time highs in march 2000 the DAX index began to fall and the downtrend lasted till the end of the observation period. From the future prediction point of view the market situation has been challenging, because there has been many sudden changes in the direction of the index. In other words, instead of unambiguous major trend there has been at least three different minor trends or market periods; the steady period of 1999, the steep rise and the downtrend in 2000-2001.



**Figure 9.** Dax index 4.1.1999 – 28.12.2001

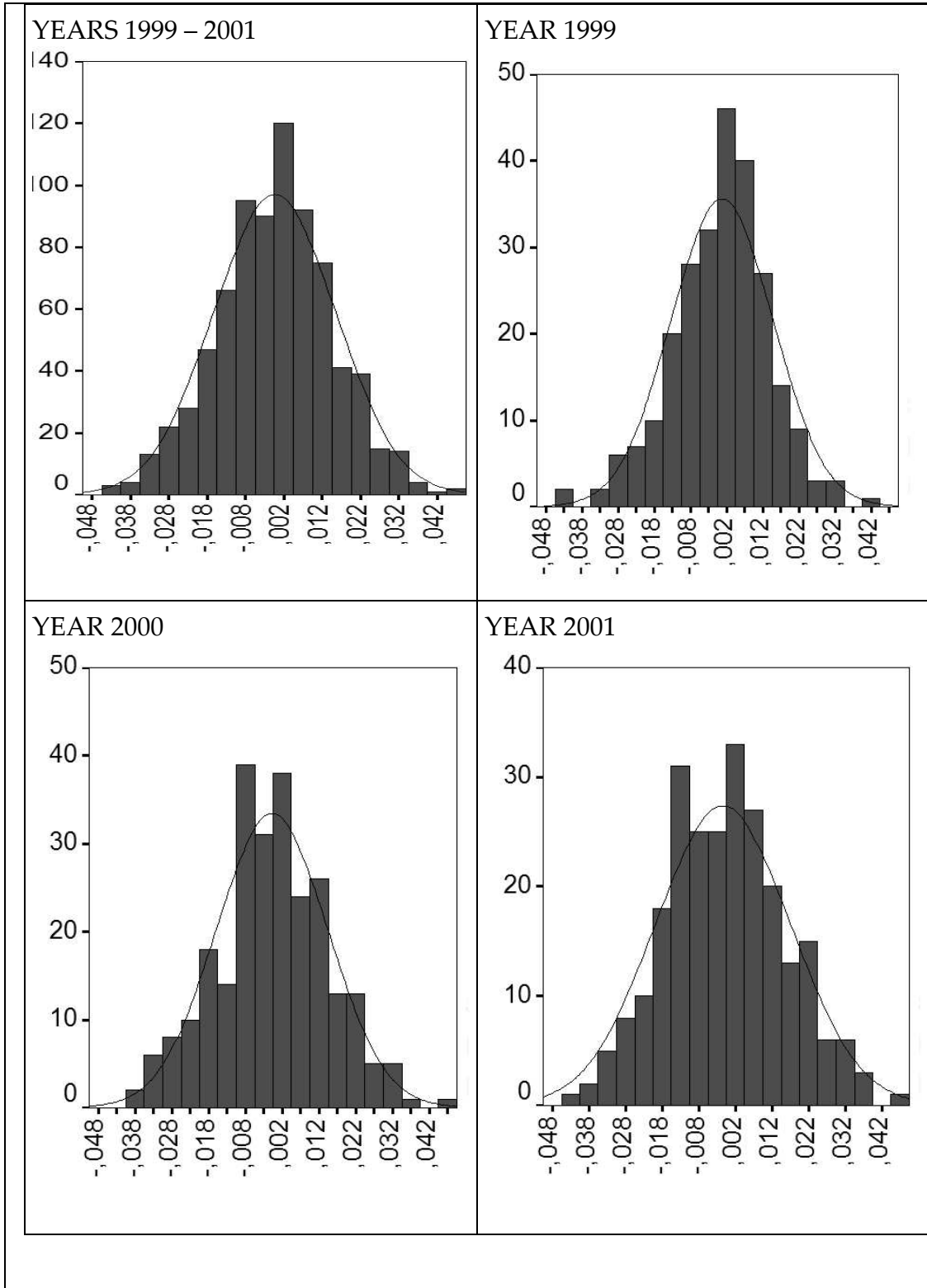
For each day the data contains settlement prices for options with at least 4 different maturities with about 20 different strike prices each. It is obvious that all the data can not be taken into the study, and it is not even the purpose. The test sample, against which the hypotheses are tested, is formed by drawing a random sample of days from the time period. The sample consists of three random draw days for each month so the total sample size is 108 days.

The characteristics of the daily returns data are graphically illustrated in figure 10, while table 1 summarises the descriptive statistics. As the descriptives clearly show, the logarithmic daily returns are not normally distributed and the entire data clearly exhibits negative skewness and excess positive kurtosis. For years 1999 and 2000 the distributions are not clearly skewed but the year 1999 exhibits non-normal kurtosis, while returns in year 2000 actually seems to be

quite normally distributed. The descriptives for year 2001 are similar as for the entire data, skewness being -0,4 and kurtosis being as high as 5,4.

**Table 1.** Statistical characteristics of Dax index daily returns for years 1999-2001

	<b>N</b>	<b>Mean</b>	<b>Median</b>	<b>Std.Dev.</b>	<b>Skewness</b>	<b>Std. Error</b>	<b>Kurtosis</b>	<b>Std. Error</b>
<b>Entire data 1999-2001</b>	759	0,00004	0,00057	0,016	-0,212	0,089	4,870	0,177
<b>1999</b>	252	0,00125	0,00140	0,014	0,071	0,153	4,414	0,306
<b>2000</b>	254	-0,00025	-0,00012	0,015	0,016	0,153	2,942	0,304
<b>2001</b>	253	-0,00087	-0,00071	0,018	-0,395	0,153	5,446	0,305



**Figure 10.** Logarithmic daily returns

For each day in the sample and for each maturity at that day, a measure for realised volatility (standard deviation), -skewness and -kurtosis to the maturity is calculated. Methods of calculating realised moments are adopted from the study by Navatte and Villa (2000). Formulas for realised volatility (RSD), skewness (RSK) and kurtosis (RKU), respectively, are the following:

$$(40) \quad RSD_i^2(T) = \frac{1}{T-1} \sum_{i=1}^T (R_{t+i} - \bar{R})^2$$

$$(41) \quad RSK_i(T) = \frac{T}{(T-1)(T-2)} \sum_{i=1}^T \left( \frac{R_{t+i} - \bar{R}}{RSD_i(T)} \right)^3$$

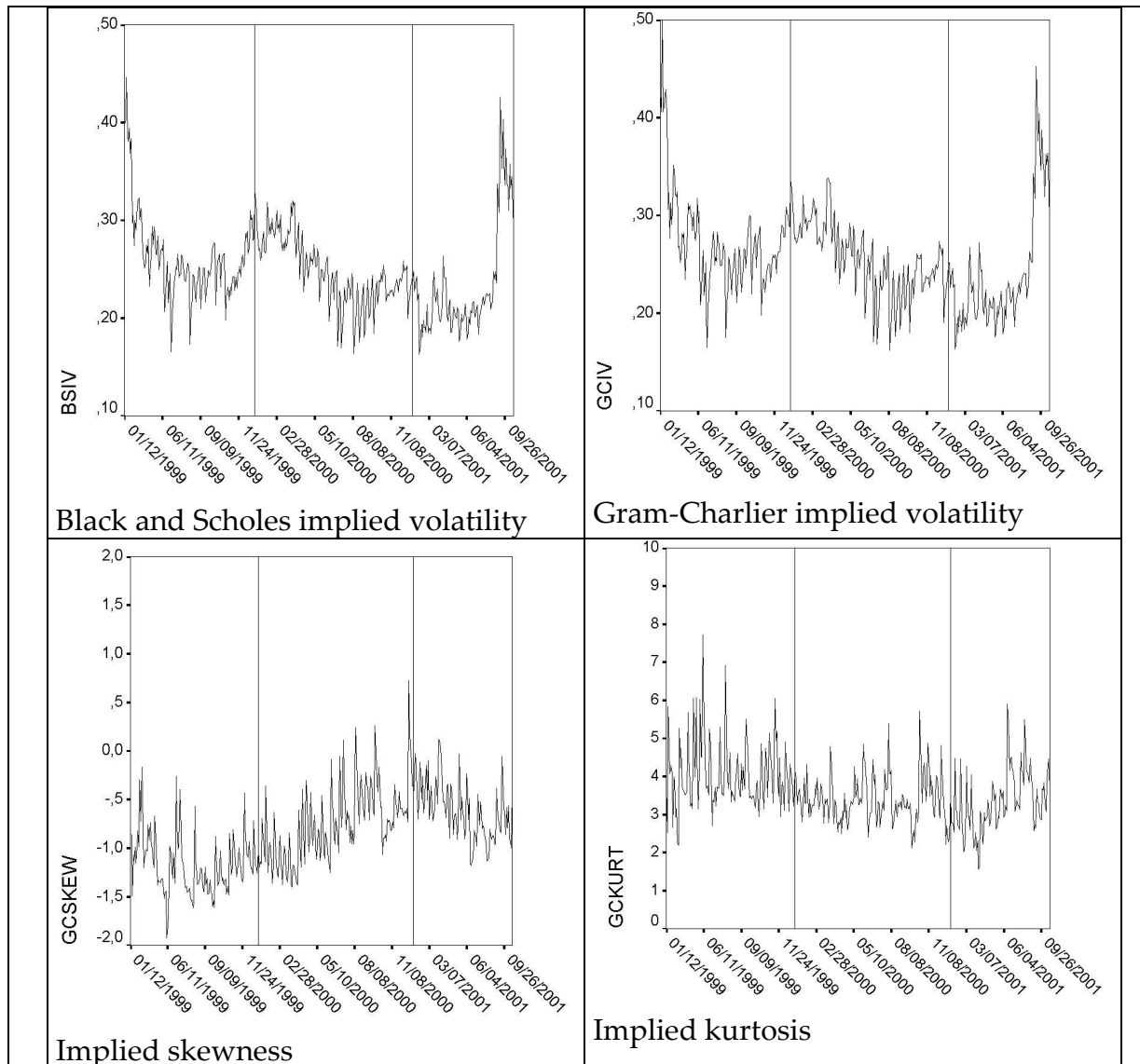
$$(42) \quad RKU_i(T) = \frac{T(T+1)}{(T-1)(T-2)(T-3)} \sum_{i=1}^T \left( \frac{R_{t+i} - \bar{R}}{RSD_i(T)} \right)^4$$

The estimation of implied volatility, skewness and kurtosis is done using least squares minimisation method, which is discussed earlier in chapter 4.2.. It is the very same method as used in studies by Corrado & Su (1996 & 1997) and Navatte & Villa (2001). Estimates for implied moments are obtained for each day and for each maturity in the data sample, so they are corresponding to historical and future realised moments. Furthermore, also the black-scholes implied volatility is calculated to enable the comparison between the prediction power of black-scholes and skewness and kurtosis adjusted models. The actual estimation procedure is carried out by using linux based GNU Octave, which is a high-level mathematical program and is mostly compatible and comparable to well known MatLab-program.

The descriptive statistics of estimates for black-scholes implied volatility, gram-charlier implied volatility, -skewness and -kurtosis are presented in table 2 below and the intertemporal variance is graphically illustrated in figure 11. As the descriptives show, on average, implied distributions are negatively skewed and exhibit positive excess kurtosis. When compared to the realised distributions, it can be seen that the implied distributions are systematically more negatively skewed. The realised skewness for entire data is -0,212 while implied skewness is on average as high as -0,854. It should be noted that neither implied volatility, skewness nor kurtosis is intertemporally stable.

**Table 2.** Descriptive statistics of the implied moments.

<b>Entire data</b> <b>1999-2001</b>	<u>Black and Scholes</u>	<u>Gram-Charlier series expansion</u>		
	Implied volatility	Implied volatility	Implied skewness	Implied kurtosis
Mean	0,24860	0,25685	-0,85371	3,59257
Median	0,24205	0,25008	-0,87155	3,46049
Std.dev.	0,04520	0,04990	0,40472	0,81977
Minimum	0,16290	0,16180	-1,92450	1,56420
Maximum	0,44620	0,50470	0,72410	7,72060
# observations	410	410	410	410
<hr/>				
<b>1999</b>	<u>Black and Scholes</u>	<u>Gram-Charlier series expansion</u>		
	Implied volatility	Implied volatility	Implied skewness	Implied kurtosis
Mean	0,26221	0,27330	-1,16935	4,03506
Median	0,25413	0,26455	-1,20530	3,81187
Std.dev.	0,04340	0,05000	0,30058	0,87602
Minimum	0,16550	0,16490	-1,92450	2,19920
Maximum	0,44620	0,50470	-0,16870	7,72060
# observations	137	137	137	137
<hr/>				
<b>2000</b>	<u>Black and Scholes</u>	<u>Gram-Charlier series expansion</u>		
	Implied volatility	Implied volatility	Implied skewness	Implied kurtosis
Mean	0,24762	0,25458	-0,75194	3,40101
Median	0,24421	0,25562	-0,79828	3,29865
Std.dev.	0,03450	0,03700	0,37645	0,61503
Minimum	0,16400	0,16180	-1,39760	2,11470
Maximum	0,33050	0,33870	0,72410	5,71250
# observations	166	166	166	166
<hr/>				
<b>2001</b>	<u>Black and Scholes</u>	<u>Gram-Charlier series expansion</u>		
	Implied volatility	Implied volatility	Implied skewness	Implied kurtosis
Mean	0,23270	0,23932	-0,60746	3,32322
Median	0,21387	0,22238	-0,63778	3,21689
Std.dev.	0,05590	0,05990	0,29881	0,79738
Minimum	0,16290	0,16340	-1,17580	1,56420
Maximum	0,42540	0,45250	0,11800	5,89350
# observations	107	107	107	107



**Figure 11.**

## 6.2. Research process

The actual testing of the research hypothesis is done with a simple regression analysis using realised moment as a response and implied moment as a predictor. Similar regression equations for realised moments and historical moments are formed. Regression equations for skewness and kurtosis are the same as equations 43 and 44 for volatility show, except the volatility being replaced by corresponding moment.

$$(43) \quad \sigma_{realised} = \alpha + \beta_1 \sigma_{historical} + \varepsilon$$

$$(44) \quad \sigma_{realised} = \alpha + \beta_1 \sigma_{implied} + \varepsilon$$

For regression analysis concerning the volatility prediction also the Black-Scholes implied volatility is used as a predictor so that the prediction power of the black-scholes implied volatility and gram-charlier implied volatility can be compared. For higher moments only the gram-charlier implied and historical measures are used, because it is not possible to obtain black-scholes implied measures for them.

Furthermore, it is explored whether historical- and implied moments contain some additional information when combined together. This is done by forming regression equations for volatility, skewness and kurtosis similar to the equation 45 below.

$$(45) \quad \sigma_{realised} = \alpha + \beta_1 \sigma_{historical} + \beta_2 \sigma_{implied} + \varepsilon$$

As already mentioned, during the observation period the market conditions have been very volatile and there has not been one distinct major trend, but many different minor trends. The market occurred steep rises, sudden collapses, hope and despair among market operators. The investment decisions are made on the basis of current market sentiment and for longer maturity options the market sentiment can change many times until the maturity. It is very difficult to predict all these changes in the market and it might be reflected on the different prediction power of long and short term implied moments. Option contracts in the sample have wide range of maturities, contract life spans ranging from 20 days to 1 year, therefore while reflecting many different investment horizons they might have different forecast power.

Because of the maturity issue, the data sample is split in two categories on the basis of maturity and then the regression equations used on the entire data earlier are applied also to this new split data. The option contracts are divided in these categories so that all options with less than or equal to 100 days to maturity are considered as short maturity options and options with more than 100 days to maturity are considered as long maturity options.

Next chapter presents and discusses the empirical results.

## 7. EMPIRICAL RESULTS

This chapter presents and discusses the empirical results. Predictive power hypotheses are tested using regression equations (43, 44 and 45) presented earlier in chapter 6 and the results for volatility, skewness and kurtosis are presented in sub-chapters below.

### 7.1. Results on implied volatility

First of all, the prediction power of historical and implied volatility on future volatility is tested using four different regression models; Historical, Implied BS, Implied GC and Historical + Implied GC. These models are presented as equations 43, 44 and 45 in previous chapter. The results for volatility prediction are presented in table 3, which summarises regression coefficients for different models. Significance levels for each coefficient are presented in parenthesis below coefficients.

The results imply that, when all maturities are examined, both, historical and implied (both BS and GC) volatilities contain a substantial amount of information for future volatility. All of those coefficients significantly differ from zero, which is quite consistent with findings of Navatte & Villa (2000) and Canina & Figlewski (1993), although Navatte & Villa (2000) report that coefficient for historical volatility does not differ from zero. Furthermore, it should be noted that the coefficients also significantly differ from the number one which is consistent with Canina & Figlewski (1993), but is in sharp contrast with Navatte & Villa (2000) who report coefficients for implied volatilities statistically close to the number one. As slope coefficients are compared it seems that coefficients of both implied volatilities (BS & GC) are slightly higher than coefficients of historical volatilities. This can be also observed from the coefficients of model combining both, historical and implied volatilities; Coefficients of implied GC volatility is slightly higher than the coefficient of historical volatility.

These four models can be compared by their  $R^2$  –measures. In this point of view it is not clear whether historical or implied volatilities should be used to predict future volatility.  $R^2$  measures for all of these four models are between 0,029 – 0,042, which can be considered to be rather low when compared to  $R^2$  measures around 0,25 which Navatte & Villa (2000) reported for implied volatility models. Still, highest  $R^2$  measure is achieved by model combining both historical and implied volatility information, so it seems like they are not exclusive and should be combined to achieve better prediction results.

When maturity issue is explored and the data is split in two categories, the results seem little different. For historical model there are not much changes, except the prediction power beyond 100 days seems to be even worse; Slope on historical volatility on maturity over 100 days does not statistically differ from zero, and  $R^2$  measure is as low as 0,002. For both models using implied volatilities, the maturity split makes a dramatic change in  $R^2$  –measures and slope coefficients. All coefficients for implied models significantly differ from both zero and the number one. For implied BS model, the  $R^2$  for all maturities is 0,034 and it improves to 0,107 and 0,163, for maturities less than or equal to 100 and over 100, respectively. Simultaneous, the slope coefficients change as the maturity increases; Slope for implied BS on all maturities is 0,248 and for maturities under or equal to 100 days it is 0,440, while being  $-0,532$  for maturities over 100 days.

So, it seems like there is a strong negative correlation between implied BS volatility and future volatility, i.e. market operators somehow overreact and fail to predict the future volatility correctly. The results for implied GC model are quite similar to implied BS model;  $R^2$  measures for implied GC models improve when maturities are split in two.  $R^2$  for all maturities is 0,029, while for being 0,095 and 0,173 for maturities under or equal to 100 and maturities over 100, respectively. Also, the slope coefficient for short maturities is 0,374 and  $-0,506$  for longer maturity options. So, there is similar change from positive to negative coefficient as there was in implied BS model. When we take a look at the model combining historical and implied volatility information, we observe similar tendency in  $R^2$  measures as in both implied models;  $R^2$  improves from 0,042 for all maturities to 0,100 for short maturity options and 0,183 for longer maturity

options. All coefficients differ significantly from the number one and all but the coefficients of historical volatilities differ significantly from zero. So it seems like implied volatility has more additional information about future volatility than the historical volatility does. Even though, coefficients of historical volatility in this combined model do not significantly differ from zero, they do contain some additional information because  $R^2$  measures are higher than for the model which contains only implied volatility information.

Implied volatility information seems to outperform historical volatility in predicting the future volatility, and thus Hypotheses  $H_1$  holds. As the maturity of options increase the  $R^2$  measures of implied volatility models seem to get better, but at the same time the coefficients turn from positive to negative. So, it seems that even though information content of implied volatilities is higher for longer maturity options, implied volatilities are not good estimate of future volatility. In other words, if implied volatility for long maturity options is relatively high, then one should expect that the future realised volatility will be relatively low, and vice versa.

**Table 3.** Regressions on volatility  
regressions on volatility

Model	Intercept	Slopes on			R <sup>2</sup>
		Historical	Implied BS	Implied GC	
<b>Historical</b>					
All maturities	0,198 *! (0,000)	0,186 *! (0,000)			0,031
Maturity <= 100	0,187 *! (0,000)	0,216 *! (0,001)			0,048
Maturity > 100	0,262 *! (0,000)	-0,072 ! (0,575)			0,002
<b>Implied BS</b>					
All maturities	0,179 *! (0,000)		0,248 *! (0,000)		0,034
Maturity <= 100	0,127 *! (0,000)		0,440 *! (0,000)		0,107
Maturity > 100	0,378 *! (0,000)		-0,532 *! (0,000)		0,163
<b>Implied GC</b>					
All maturities	0,187 *! (0,000)			0,207 *! (0,001)	0,029
Maturity <= 100	0,141 *! (0,000)			0,374 *! (0,000)	0,095
Maturity > 100	0,376 *! (0,000)			-0,506 *! (0,000)	0,173
<b>Historical + GC Implied</b>					
All maturities	0,174 *! (0,000)	0,134 *! (0,018)		0,141 *! (0,031)	0,042
Maturity <= 100	0,135 *! (0,000)	0,081 ! (0,245)		0,323 *! (0,000)	0,100
Maturity > 100	0,416 *! (0,000)	-0,163 ! (0,168)		-0,522 *! (0,000)	0,183

\* Significantly different from zero at the 95% confidence level

! Significantly different from one at the 95% confidence level

## 7.2. Results on skewness

The prediction power of historical and implied skewness on future skewness is tested using three different regression models; Historical, Implied GC and Historical + Implied GC. These models are presented as equations 43, 44 and 45 in chapter 6. The results for skewness prediction are presented in table 4, which summarises regression coefficients for different models. Significance levels for each coefficient are presented in parenthesis below coefficients.

The results show that when all maturities are examined, historical skewness seems to contain some information about future skewness, because the correlation coefficient significantly differs from zero and the number one, whereas the coefficient for implied GC skewness does not significantly differ from zero, meaning that implied GC skewness does not seem to contain any additional information about future skewness. The  $R^2$  –measure for historical model is 0.022 and for implied model it is 0.000, which also means that implied skewness can not explain future skewness, as long as entire maturity range is concerned. The regression coefficients on combined model also indicate that historical skewness has more information content about future skewness than the implied skewness does. Slope on historical skewness significantly differs from zero and the number one, while slope on implied skewness does not significantly differ from zero, and thus has low information content.

Splitting data in short and long maturity options reveals similar phenomenon as in the case of volatility. For historical skewness coefficient does not significantly differ from zero for short maturity options, but for longer maturity options the slope changes to  $-0.708$ , which significantly differs from zero. At the same time  $R^2$  –measures, being 0.022 for all maturities, change to 0.014 and 0.077 for short and longer maturity options, respectively. This implies that historical skewness contains some information about longer term future skewness, but as  $R^2$  measure is so low it may be difficult to implement in prediction model. As we take a look at the implied skewness model, splitting the data in two maturity categories significantly improves prediction power of the model. For all maturities, the coefficient for implied skewness was not significantly different from zero, but after the maturity split, slopes for both,

short and long maturity options, significantly differ from zero and thus contain additional information about future skewness. Similarly as in the case of volatility, slope for short maturity options is positive and for long maturity options it is negative. Also the  $R^2$  measures improve significantly due to maturity split, and it seems like implied skewness model outperforms historical model. Furthermore, as combined model yields highest  $R^2$  measures for both short and long maturity options and coefficients significantly differ from zero, it seems like historical and implied skewness do have additional information when combined together. But still, it is not clear which of these, historical or implied skewness, is contains more information about the future skewness. For short maturity options, implied GC skewness has steeper slope than the historical, but for longer maturity options, historical skewness has a little more steeper slope than what implied skewness does.

For entire maturity range, it seems that historical skewness contains more information about future skewness, and thus hypotheses  $H_2$  fails. But when the data is split in two maturity categories, the implied GC skewness models seems to perform a little better in predicting future skewness. As we investigate the model combining both historical and implied skewness, it is uncertain which contains more information, but it is evident that both historical and implied skewness should be included in model to achieve better prediction power.

**Table 4.** Regressions on skewness  
regressions on skewness

Model	Intercept	Slopes on		R <sup>2</sup>
		Historical	Implied GC	
<b>Historical</b>				
All maturities	-0,112 *! (0,000)	-0,187 *! (0,002)		0,022
Maturity <= 100	-0,038 ! (0,203)	-0,122 ! (0,063)		0,014
Maturity > 100	-0,250 *! (0,000)	-0,708 *! (0,000)		0,077
<b>Implied GC</b>				
All maturities	-0,084 ! (0,108)		0,015 ! (0,848)	0,000
Maturity <= 100	0,093 ! (0,120)		0,153 *! (0,028)	0,019
Maturity > 100	-0,642 *! (0,000)		-0,455 *! (0,000)	0,144
<b>Historical + GC Implied</b>				
All maturities	-0,091 ! (0,081)	-0,19 *! (0,002)	0,026 ! (0,642)	0,023
Maturity <= 100	0,087 ! (0,152)	-0,134 *! (0,041)	0,163 *! (0,019)	0,036
Maturity > 100	-0,612 *! (0,000)	-0,459 *! (0,018)	-0,391 *! (0,000)	0,174

\* Significantly different from zero at the 95% confidence level

! Significantly different from one at the 95% confidence level

### 7.3. Results on kurtosis

The prediction power of historical and implied kurtosis on future kurtosis is tested using three different regression models; Historical, Implied GC and Historical + Implied GC. These models are presented as equations 43, 44 and 45 in chapter 6. The results for kurtosis prediction are presented in table 5, which

summarises regression coefficients for different models. Significance levels for each coefficient are presented in parenthesis below coefficients.

When all maturities are examined, the results on kurtosis are quite opposite to results on skewness; Historical kurtosis does contain information about future kurtosis, whereas implied GC kurtosis does. When the correlation coefficients of different models are compared, it seems that the slope on historical kurtosis does not significantly differ from zero, but the slope on implied GC is significantly different from zero, these findings mean that implied GC kurtosis has higher information content about future kurtosis than historical kurtosis. The  $R^2$ -measures for historical model is 0.001 and for implied model it is 0.010, although neither of these is very high or even at decent levels it is evident that the implied model is better than the historical model in predicting the future kurtosis. The regression coefficients on combined model similarly indicate that implied GC kurtosis contains more information about future kurtosis than the historical kurtosis. It should also be noted that regression coefficients for both historical and implied model are negative, so neither historical or implied kurtosis is a good unbiased estimate of future kurtosis.

When the data is split in short and long maturity options, the coefficients and  $R^2$ -measures change dramatically. Results for historical kurtosis model are the following: for short maturity options the regression coefficient does not significantly differ from zero and thus does not contain any significant information, but the regression coefficient for longer maturity options is  $-0.919$ , which is significantly different from zero. At the same time  $R^2$ -measures, being 0.001 for all maturities, change to 0.000 and 0.056 for short and longer maturity options, respectively. These statistics imply that historical kurtosis does contain important information about long term future kurtosis, but fails to predict short term future kurtosis.

As we focus on the implied GC model, splitting the data in two maturity categories slightly improves the prediction power of the model. For all maturities and for both split maturities the regression coefficients stay significantly different from zero, and the  $R^2$ -measures improve from 0.010 for all maturities to 0.017 and 0.047 for short and long maturity options,

respectively. Furthermore, as for results on volatility and skewness the regression coefficients for short and long maturity options have different signs, coefficient for short maturity being positive and for long maturity it is negative.

Similarly as the results on skewness, the combined model yields higher R<sup>2</sup>-measures for all maturities and both split maturities than historical or implied model. So it seems that historical and implied model do have some additional information when combined together. Actually, historical kurtosis seems to improve the prediction power of the model only for long maturity options, because historical kurtosis regression coefficients for all maturities and short maturities does not significantly differ from zero and thus only implied kurtosis seems to have any significance for all maturities and for short maturity options. But for longer maturity options, coefficients on both historical and implied kurtosis differ significantly from zero, and historical kurtosis seems to have even steeper slope than implied kurtosis, which implies that historical kurtosis is more information rich than implied kurtosis.

For entire maturity range, it seems that implied kurtosis contains more information about future kurtosis than historical kurtosis, and thus hypotheses H<sub>3</sub> holds. But when the data is split in two maturity categories, the historical kurtosis model seems to perform a little better in predicting long term future kurtosis and the implied kurtosis seems to perform better in predicting short term future kurtosis. It should be noted that R<sup>2</sup>-measures for kurtosis prediction models are much lower than measures for volatility or skewness prediction models, so even though implied and historical kurtosis contain some information about future kurtosis these models obviously need some other inputs which would have additional information about future kurtosis.

**Table 5.** Regressions on kurtosis  
regressions on kurtosis

Model	Intercept	Slopes on		R <sup>2</sup>
		Historical	Implied GC	
<b>Historical</b>				
All maturities	0,392 *! (0,000)	-0,049 ! (0,490)		0,001
Maturity <= 100	0,055 ! (0,344)	0,013 ! (0,807)		0,000
Maturity > 100	1,066 * (0,000)	-0,919 *! (0,002)		0,056
<b>Implied GC</b>				
All maturities	0,475 *! (0,000)		-0,154 *! (0,046)	0,010
Maturity <= 100	-0,045 ! (0,545)		0,131 *! (0,038)	0,017
Maturity > 100	1,072 * (0,000)		-0,622 * (0,006)	0,047
<b>Historical + GC Implied</b>				
All maturities	0,482 *! (0,000)	-0,048 ! (0,502)	-0,153 *! (0,047)	0,011
Maturity <= 100	-0,047 ! (0,535)	0,012 ! (0,823)	0,130 *! (0,039)	0,018
Maturity > 100	1,222 * (0,000)	-0,849 *! (0,004)	-0,564 *! (0,011)	0,094

\* Significantly different from zero at the 95% confidence level

! Significantly different from one at the 95% confidence level

## 8. CONCLUSIONS

The purpose of this study was to investigate the information content of option implied probability distribution about future price behaviour. This investigation was done by using volatility, skewness and kurtosis as a proxy of the underlying distribution. Furthermore, these implied and historical moments were used in regression models and regressed against future realised volatility, skewness and kurtosis.

Earlier research related to the topic can be divided in three categories, Implied Volatility, Estimation of Implied Distribution and Information Content of Implied Distributions.

Studies on implied volatility indicate that implied volatility does contain some information about future volatility and in many cases it outperforms historical volatility in predicting future volatility. It also seems that implied volatility alone does not adequately represent investors' beliefs of future price behaviour, because it measures only uncertainty, not the direction of the market. This is why we are interested in other option implied information such as implied distributions, which could reveal much more information than just the volatility.

Shape of underlying return distribution reflects the probable market direction and probability of extreme events affecting stock price. The shape of the distribution can be characterised by standard deviation (volatility), skewness and kurtosis. Volatility can be easily estimated from market prices, but obtaining skewness and kurtosis is a more complex task. Earlier studies on estimating implied distributions present few different methods how to do it, and a method by Corrado and Su (1996), which simultaneously estimates volatility, skewness and kurtosis, is selected to this study.

Information content of implied distributions and higher moments of distribution is also studied in few earlier papers. Results are contradictory, some studies suggesting that implied distributions contain significant amount of information about future realized distribution while others find quite

opposite results. Earlier studies also show that even though implied moments contain some additional information about future distribution, the prediction power of skewness and kurtosis on future skewness and kurtosis is poor.

The results of the current study are parallel to earlier studies. Overall results show that both historical and implied moments contain some information about corresponding future moments, which are volatility, skewness and kurtosis, but it is not clear whether implied information outperforms historical information. When the whole maturity range is examined, it seems that for volatility and kurtosis the implied information outperforms historical information in predicting the future, but for skewness the results are opposite; implied skewness seems to have no information about future skewness, while historical skewness has. In every case for all maturities, it seems that combining both historical and implied information improves the prediction power of model. So, both historical and implied moments should be used together in model building to achieve better prediction power.

The results show also that the prediction power of models using these moments seem to deteriorate with higher moments. Prediction power is measured by  $R^2$  and models predicting volatility get highest  $R^2$ -measures, while it is a little lower for skewness and lowest for kurtosis prediction models. Furthermore, when the data is split in two maturity categories, prediction power of all these models improves dramatically. In general, it seems that these models perform better in predicting long term price behaviour and the long term, meaning maturity greater than 100 days, combined model yields highest  $R^2$ -measures for all these moments, volatility, skewness and kurtosis.

The regression coefficient for the short and long maturity models usually have opposite signs. This phenomenon and the results on prediction power indicate that the correlation coefficients of historical and implied models are not constant and thus it is beneficial to use different models to different time horizons.

Overall, it seems that implied volatility, skewness and kurtosis do contain some information about the future volatility, skewness and kurtosis, but as the

prediction power of these models used in this study is so low, it is difficult to implement this information on predicting the future.

## REFERENCES

- Abken, P., D. Madan, S. Ramamurtic (1996). Estimation of Risk Neutral Densities by Hermite Polynomial Approximation: With an Application to Eurodollar Future Options. Federal Reserve Bank of Atlanta, *Working paper 96-5*.
- Anagnou, I., M. Bedendo, S. Hodges, R. Tompkins (2002). The relation between implied and realised probability density functions. DRAFT. Available on the internet
- Beckers, S. (1981). Standard deviations implied in option prices as predictors of future stock price variability. *Journal of Banking and Finance* 5, 363-382.
- Breedon, D., R. Litzenberger (1978). Prices of State-Contingent Claims Implicit in Option Prices. *Journal of Business* 9, 621-651.
- Black, F. & M. Scholes (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* 81, 637-654.
- Bates, D. (1996) Jumps and Stochastic volatility: Exchange rate process implicit in deutchemark options. *Review of Financial Studies* 9, 69-108.
- Cox, J., M. Rubinstein(1985). *Option Markets*. Prentice Hall, Inc., New Jersey.
- Cox, J., S. Ross (1976). The Valuation of options for alternative stochastic processes. *Journal of Financial Economics* 3, 145-166.
- Cox, J., S. Ross & M. Rubinstein (1979). Option Pricing: A simplified approach. *Journal of Financial Economics* 7, 229-263.
- Corrado, C., T. Su (1996). Skewness and kurtosis in S&P 500 Index returns implied by option prices. *Journal of Financial Research*. Vol. XIX, no.2. 175-192.

- Corrado, C., T. Su (1997). Implied volatility skews and stock return skewness and kurtosis implied by stock option prices. *European Journal of Finance* 3, 73-85.
- Churchill, R. (1963) *Fourier Series and Boundary Value Problems*. 2<sup>nd</sup> ed. New York: McGraw-Hill.
- Deutsche Börse (2002). Short info on DAX. Available on internet <URL:[http://www.ip.exchange.de/INTERNET/IP/ip\\_stats.nsf/WebMaskskenformeln/20F6D83A810FC35541256A37004BB077/\\$file/Dax\\_i\\_e.pdf](http://www.ip.exchange.de/INTERNET/IP/ip_stats.nsf/WebMaskskenformeln/20F6D83A810FC35541256A37004BB077/$file/Dax_i_e.pdf)>
- Day, T., C. Lewis (1992) Stock market volatility and the information content of stock index options. *Journal of Econometrics* 52, 267-287.
- Canina, L., S. Figlewski (1993). The informational content of implied volatility. *Review of Financial Studies* 6:3, 659-681.
- Dumas, B., J. Fleming & R. Whaley (1998). Implied Volatility Functions: Empirical tests. *The Journal of Finance* 53, 2059-2106.
- Dennis, P. & S. Mayhew (2002). Risk-Neutral Skewness: Evidence from Stock Options. *Journal of Financial and Quantitative Analysis*, Vol. 37, No. 3, 471-494.
- Fornari, F. & A. Mele (2001). Volatility Smiles and the Information Content of News. *Applied Financial Economics* 11, 179-186
- Hull, J., A. White (1987). The Pricing of Options on Assets with Stochastic Volatilities. *Journal of Finance*. Vol. 42. 281-300.
- Hull (2000). *Options, Futures and Other Derivatives*. Upper Saddle River (N.J.) : Prentice-Hall, cop. 2000

- Ito, K. (1951). On Stochastic Differential Equations. *Memoirs, American Mathematical Society* 4, 1-51.
- Jackwerth, J., M. Rubinstein (1996) Recovering Probability Distributions from Option Prices. *Journal of Finance*. Vol 51, no.5. 1611-1631.
- Jarrow, R., A. Rudd (1982) Approximate Option Valuation for Arbitrary Stochastic Processes. *Journal of Financial Economics*, 10 347-369.
- Jorion, P. (1995). Predicting Volatility in the Foreign Exchange Market. *Journal of Finance* 50:2, 507-528.
- Latané, H., R. Rendelman. (1976). Standard deviation of stock price ratios implied by option premia. *Journal of Finance*, 31 368-382.
- Longstaff, F. (1990). Option Pricing and the Martingale Restriction. *Review of Financial Studies* 8:4, 1091-1124.
- Madan, D., F. Milne (1994) Contingent Claims valued and Hedged by Pricing and Investing in a Basis. *Mathematical Finance*, 4 223-245.
- Merton, R. (1976). Option Pricing When Underlying Stock Returns Are Discontinuous. *Journal of Financial Economics*. Vol. 3. 125-144.
- Merton, R. (1973) The theory of rational option pricing. *Bell Journal of Economics and Management Science*. Vol. 4. 141-83
- Nikkinen (2001). Skewness and kurtosis adjusted Black-Scholes model: Small sample properties of MLE and test statistics and empirical evidence from Finland. *Acta Wasaensia*. No. 94
- Navatte, P., C. Villa (2000). The information content of implied volatility, skewness and kurtosis: Empirical evidence from long-term CAC 40 options. *European Financial Management* 6, 41-56.

- Rubinstein, M. (1994). Implied binomial trees. *Journal of Finance* 69, 771-818.
- Smith, B., F. Merceret (2000). The Lognormal Distribution. *The College Mathematics Journal*. Vol. 31. No. 4. 259-261.
- Shiratsuka, S. (2001) Information content of implied probability distributions: empirical studies on japanese stock spice index options. *Monetary and Economic Studies*.
- Shimko, D. (1993). Bounds of Probability. *Risk* 6:4, 33-37.
- Weinberg, S.A. (2001). Interpreting the Volatility Smile: An Examination of the Information Content of Option Prices. International Finance Discussion Papers. Available on internet <URL: <http://www.federalreserve.gov/pubs/ifdp/>>
- Whaley, R. (1982) Valuation of American Call Options on Dividend Paying Stocks. *Journal of Financial Economics*, 10:29-58.

