Nationalekonomiska Institutionen

Bachelor's thesis
January 2005

The Relation between information in option prices and short term market return

Instructor
Birger Nilsson

Student
Henrik Stille

Abstract

The purpose of this thesis is to analyze if implied volatility and the implied risk neutral density function can predict short term market return of the Swedish OMX-index. To study the relation between information in option prices and short term market return I perform and OLS-regression with the market return as dependent variable. Implied volatility and the various moments of the implied density function are used as independent variables.

Data from the Swedish OMX-index at the Stockholm Stock Exchange for the period September through December 2002 is used in the study.

The study shows that high implied volatility of at the money options predicts a positive return on the OMX-index in the next three days. The prediction power is also strong for the third moment of the implied density function i.e. skewness.

Key words: option prices, implied volatility, short term return, mean reversion, implied risk neutral density function

Table of Contents

1 Introduction

1.1 Background

The most valuable ability a person that works in the financial industry can have is an ability to be able to forecast the markets. It is through superior forecasting ability that it is possible to make money in the market. Enormous amounts of money have been spent on research over the last 50 years to develop models that can be used for forecasting. Markets can be forecasted through either fundamental or technical analysis. A fundamental analyst believes that all historical information already is priced in the market i.e. the market is weakly efficient. The fundamental analyst does therefore not spend any time on analyzing charts or other historical information; instead he focuses on the market fundamentals and tries to find markets that are under-priced when considering the analyst's beliefs of the future.

The technical analyst does not believe that the market is weakly efficient. He therefore thinks it is worthwhile to study historical information and historical charts and through this find markets that will perform better than the average market. The technical analyst may look at long term trends, volume behavior or other indicators that makes the analyst believe that a certain pattern will repeat in the future. The idea behind the technical analysis is that observable factors other than the stock price itself convey information about future movements e.g. movements in volume or long term trends make it likely that the stock in the future will move in a particular way.

Many private investors find technical analysis appealing since the information needed to perform these analysis is easy to obtain, it is much harder for the individual investor to obtain fundamental information about companies and even harder to use fundamental information to make a reliable forecast.

Derivatives are another source of information that can be used to predict upcoming market movements. The price of a future contract that expires some time in the future should equal the current spot price of the underlying asset plus the time value of money. It is very common that this relation does not hold. Deviations from this relation are due to the fact that the market expectation about the price of the asset in the future is different from the price today. This reasoning suggests that prices of futures and forwards conveys information about future movements in the underlying asset. We will see below that the same reasoning can be applied to options.

In the last two decades the use of derivatives for different purposes has increased dramatically around the world. Especially in the foreign exchange market various derivatives have become more common e.g. the turnover of currency forwards and futures contracts increased more than ten times during the 1990's.[1] The same pattern is true for stock options that today are traded in a multiplicity of maturities and strike prices with various issuers. This is a characteristic of a well-functioning and efficient secondary market.[2] The fact that the market

[1] David K. Eiteman, Arthur I. Stonehill, Michael H. Moffett, Multinational Business Finance (2004) p78.
[2] Lars Oxelheim, Financial Markets in Transition (1996) p7 ff.

for futures and options today are much more efficient and well functioning than 20 years ago makes them attractive for more investors and hence option prices reflect the view of many more investors today than they did 20 years ago. This makes derivatives more interesting for forecasting.

1.2 Problem discussion

The interest for options as a forecasting instrument took a big turn up in the early 1990's when Mark Rubinstein published his theory about crashophobia. According to Rubinstein traders are so concerned about the possibility of another crash similar to October 1987 that they price options with a pronounced skew[3] i.e. the implied volatility of stock options are much lower for options with high strike prices than for options with low strike prices i.e. traders pay a higher price for deep out of the money put options than deep out of the money call options.

There are several explanations for crashophobia. One possibility is that negative shocks to returns drive up volatility. The leverage hypothesis, due originally to Black (1976)[4] says that when the total value of a levered firm falls the value of its equity becomes a smaller share of the total. Since equity bears the full risk of the firm the percentage volatility of equity should rise. Even if a firm is not financially levered with debt this may occur if the firm has fixed commitments to workers or suppliers.

Jackwerth & Rubinstein (2001)[5] argues that equity prices have become more highly correlated in down markets causing an increase in volatility. Secondly, risk aversion effects can cause investors who are poorer after a downturn in the market to react more dramatically to news events. This would lead to increased volatility after a downturn. Thirdly, the market in general could be more likely to jump down rather than up since the five greatest moves in the S&P 500 index after the crash of 1987 have been down. Finally as the volatility of the market increases the required risk premium rises too and a higher risk premium will in turn depress stock prices. I will return to the concept of implied volatility and volatility skewness later in this study.

How is it then possible to use the information in option pricing to forecast the market? According to Mark Rubinstein the volatility skewness becomes more pronounced in bear markets and less pronounced in bull markets. But what does this tell us about future movements? What Mark Rubinstein says is really only that the skew is more pronounced in bear markets, this does not necessarily tell us anything about if the bear market will continue, or does it? In this study I will argue that a heavy pronounced skew is a signal of a future upturn in the market, at least in the short-term. The logic behind this argument is that when the skew is heavily pronounced the market participants are very pessimistic about future returns. This is due to a combination of the above mentioned factors. Due to the mean-reversion phenomenon such irrational pessimism will make the market turnaround, at least in the short term since according to the mean reversion phenomenon the market in the short run

[3] John C. Hull, Options, Futures and Other Derivatives (2003) p335.
[4] Campbell, Lo, McKinlay, The Econometrics of Financial Markets (1997) p497.
[5] Jens Carsten Jackwerth and Mark Rubinstein, Recovering Stochastic Processes from Option Prices The Journal of Finance (2002) p 1611.

is expected to return to its long term mean[6]. Strong positive volatility skewness should therefore be positively related to upcoming short term return. The evidence of mean-reversion in stock markets is strong and in this study I will use the information in option prices to measure how close the market is to revert to its long term mean.

The volatility skew will in this study be analyzed through the so called implied risk neutral density function. There exist an exact relation between the volatility skew and the implied density function.[7] I will use the implied density function since it then also will be possible to study other moments of the implied distribution than skewness. I will in addition to study the skewness also study the mean, the volatility and kurtosis of the implied distribution.

Another approach is to study how the implied volatility emerges from day to day for a certain strike price without considering the relation of implied volatility between different strike prices. One could argue that high implied volatility occurs when the market expects a large movement in either direction, and low implied volatility occurs when the market expects only small movements. One can though also argue that this is not of much help though since we get no information about if the market will move up or down.

However, in this study I will argue that increasing and high implied volatility is much more associated with nervousness about future downturns than euphoria about future upturns. I base this argument on research in the stock market that shows that there appears to be an asymmetry in stock market data. As French, Schwert and Stambaugh (1987) emphasize there is much stronger evidence that positive innovations to volatility are correlated with negative innovations to returns than vice versa.[8] An alternative explanation is that causality runs the other way: Positive shocks to volatility drive down returns. Campbell and Hentschel (1992) call this the volatility feedback hypothesis. If expected stock returns increase when volatility increases and if expected dividends are unchanged then stock prices should fall when volatility increases.[9]

Even if these arguments not are completely transferable to implied volatility in options I still think that it is a solid base for my argument that higher implied volatility in options are correlated with bear markets i.e. option premiums rise in bear markets due to the fact that writers of options demand a higher premium compare to in bull markets. The second part of this study will be to analyze the relation between implied volatility and future market return.

If we assume that it is true that high implied volatility is correlated with downturns how can we then use this information to make forecasts about the future? The explanation is the same as for the volatility skewness above. Implied volatility shows over time a typical mean reversion pattern i.e. high implied volatility is followed by lower implied volatility and consequently lower option premiums and positive market return. When the implied volatility is abnormally high the mean-reversion phenomenon should lead to an upturn in the market, at least in the short term.

[6] Ronald Balvers, Yangru Wu, Erik Gilliland, Mean reversion across National Stock Markets and Parametric Contrarian Investment Strategies, The Journal of Finance (2000) No. 2.
[7] John C. Hull, Options, Futures and Other Derivatives (2003) p334.
[8] Campbell, Lo, McKinlay, The Econometrics of Financial Markets (1997) p497.
[9] Campbell, Lo, McKinlay, The Econometrics of Financial Markets (1997) p497.

Compared to other studies in this area I will focus this study entirely on short term relations between the return in the OMX-index and 1) implied volatility and 2) various moments of the implied density function. The purpose of this study will therefore be to analyze how information in option prices revealed in form of implied volatility and the implied density function is related to upcoming short-term return in the OMX-index.

The study is limited to OMX-index options on the Stockholm Stock Exchange. Further on I will only study short term relations i.e. relations of between one and three days. There will be no attempts to analyze or speculate in any relations beyond the three day periods. All relations that are discussed in this study are short term relations. This might seem strange, but a large amount of the trades in OMX-options are conducted with a very short time horizon in mind. Due to the time consuming process of calculating the implied density functions the study is also limited in time. The calculations are done in Excel and to be able to study a longer period of time a more advanced program (e.g. MatLab) would be preferable.

The rest of this thesis is organized as follows. In chapter two I will present some general option theory, the Black-Scholes model and theory about implied volatility. In chapter three follows a survey about the used methodology and I also present the hypotheses that are used in this study. In chapter four I present the empirical results of the study and the thesis end in chapter five with conclusions and suggestions for further research.

2 Theory

In this chapter I will go through the theory that I am basing my arguments in this study upon. I start with some general theory about options, then follows theory about implied volatility and the implied density function respectively. I end the chapter with small selections from related prior research.

2.1 General option theory

Options are today issued on various underlying assets; stocks, bonds, currencies and commodities to name a few, and new areas where options are useful are rapidly being added. In the state of Florida in the United States options on the weather currently have been issued Weather options are supposed to be popular among farmers that want to insure themselves against abnormal weather that could affect the crops.[10] In this study I will focus on index options and if nothing else is stated in the text it is index options that I am referring to when I mention options.

There are two basic types of options; calls and puts. A call option gives the holder the right but not the obligation to buy the underlying asset by a certain date for a certain price. A put option gives the holder the right but not the obligation to sell the underlying asset by a certain date for a certain price. Both calls and puts exist in two varieties; American and European. American options can be exercised at any time up to the expiration date and European options can only be exercised at the expiration date.

There are six factors that affect the price of a stock option:[11]

1. The current stock price, S_0
2. The strike price, K
3. The time to expiration, T
4. The volatility of the stock price, σ
5. The risk free interest rate, r
6. The dividends expected during the life of the option

This study will focus entirely on the volatility, for those interested in how the other factors affect the price of a stock option I recommend Hull (1997).

The volatility of a stock price is a measure of how uncertain we are about future stock price movements; when volatility increases the chance that the stock will do very well or very poorly increases. For the owner of a stock these two outcomes tend to offset each other but this is not the case for an owner of a call or a put. The owner of a call benefits from price increases but has limited downside risk in the event of price decreases because the most the owner can loose is the price of the option. Similarly the owner of a put benefits from price

[10] CNBC Business news 10/27/2004.
[11] John C. Hull, Options, Futures and Other Derivatives (2003) p167.

decreases but has limited downside risk in the event of price increases. The value of both calls and puts therefore increases as volatility increases.[12]

2.2 The Black-Scholes model

In the early 1970's Fisher Black, Myron Scholes and Robert Merton made a major breakthrough in the pricing of stock options. This involved the development of what has become known as the Black-Scholes model. The model has heavily influenced the way traders price and hedge options.[13] In 1997 the importance of the model was recognized when Robert Merton and Myron Scholes were awarded the Nobel Prize in economics.

The key assumption behind the Black-Scholes model is that stock prices follow a geometric Brownian motion. A stock price that follows a geometric Brownian motion has a constant expected return.

$$\delta S = \mu S \delta t \qquad \textit{Equation 2.1}$$

where S is the stock price at time t and the parameter μ is the expected rate of return on the stock. This means that in a short interval of time the expected increase in S is $\mu S\delta t$.

In practice stock prices exhibit volatility. A reasonable assumption is that the variability of the percentage returns in a short period of time δt is the same regardless of the stock price. This suggests that the standard deviation of the change in a short period of time δt should be proportional to the stock price and leads to the model:

$$\delta S = \mu S \delta t + \sigma S \varepsilon \sqrt{\delta t} \qquad \textit{Equation 2.2}$$

where σ is the volatility of the stock price and ε is a random drawing from a standardized normal distribution i.e. a normal distribution with a mean of zero and standard deviation of 1. Both σ and μ are assumed to be constant. By expressing equation 2.2 in continues time ($\delta t \rightarrow 0$) and by dividing it with S the stock price dynamics becomes:

$$\frac{dS}{S} = \mu dt + \sigma dz \qquad \textit{Equation 2.3}$$

where dz is a continuous Wiener process equal to $\varepsilon\sqrt{dt}$. Applying a mathematical result known as Ito's lemma[14] to equation 2.3 it can be shown that the price f of an option or another derivative written on the underlying stock S, has to satisfy the following relation:

$$df = \left(\frac{df}{dS}\mu S + \frac{df}{dt} + \frac{1}{2}\frac{d^2 f}{dS^2}\sigma^2 S^2\right) dt + \frac{df}{dS}\sigma S dz \qquad \textit{Equation 2.4}$$

[12] John C. Hull, Options, Futures and Other Derivatives (2003) p168.
[13] John C. Hull, Options, Futures and Other Derivatives (2003) p234.
[14] For details about Ito's lemma please see Hull (1997) p226.

By comparing equation 2.3 and 2.4 we see that both *S* and *f* are affected by the same source of uncertainty, *dz*. This is a very important result in the derivation of the Black-Scholes model. Equation 2.4 is the starting point for deriving the famous Black-Scholes-Merton differential equation. By combining the stock and the derivative in the same portfolio the stochastic component $dz = \varepsilon\sqrt{dt}$ can be eliminated, making the portfolio riskless. This is always possible since the stock and its derivative are affected by the same source of risk.[15] The riskless portfolio can be obtained by going short in one derivative and long in an amount of $\frac{df}{dS}$ shares. To eliminate the possibility for arbitrageurs to make riskless profit this portfolio must instantaneously earn the same rate of return as other short-term securities i.e. the return on this portfolio must equal the risk free interest rate, *r*. This gives us the Black-Scholes-Merton differential equation:

$$rf = \frac{df}{dt} + rS\frac{df}{dS} + \frac{1}{2}\sigma^2 S^2 \frac{d^2 f}{dS^2} \qquad \textit{Equation 2.5}$$

It is important to realize that the portfolio used to derive equation 2.5 only is riskless instantaneously. When *S* and *t* change also df/dS will change. For the portfolio to stay riskless it must be continuously rebalanced.[16] The Black-Scholes-Merton differential equation can be used to find the price of many different types of derivatives with the price *S* of a non-dividend paying stock as the underlying variable by applying the relevant condition on the boundary value, e.g. C(T)=max[S(T)-K,0] for a European call option. The Black-Scholes formula for the price at time zero of a European call option on a non-dividend paying stock becomes:

$$c = S_0N(d_1) - Ke^{-rT}N(d_2) \qquad \textit{Equation 2.6}$$

where

$$d_1 = \frac{\ln(S_0/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}} \qquad \textit{Equation 2.7}$$

$$d_2 = d_1 - \sigma\sqrt{T} \qquad \textit{Equation 2.8}$$

Solving the differential equation above is one way of deriving the Black-Scholes formula. Another approach is to use risk neutral valuation. It arises from one key property of the Black-Scholes-Merton differential equation 2.5. This property is that the equation does not involve any variable that is affected by the risk preferences of investors. The variables that appear in the equation are the current stock price, time, price volatility and the risk free interest rate. All of these are independent of risk preferences.

The Black-Scholes-Meron differential equation 2.5 would not be independent of risk preferences if it involved the expected return on the stock, μ. It is therefore fortunate that μ

[15] Stig Arild Syrdal, A study of Implied Risk Neutral Density Functions in the Norwegian Option Market, Working paper, Norges Bank (2002) p6.
[16] Stig Arild Syrdal, A study of Implied Risk Neutral Density Functions in the Norwegian Option Market, Working paper, Norges Bank (2002) p7.

happens to drop out in the derivation of the differential equation. Because the Black-Scholes-Merton differential equation is independent of risk preferences an ingenious argument can be used. If risk preferences do not enter the equation they cannot affect its solution. Any set of risk preferences can therefore be used when evaluating *f*. In particular the very simple assumption that all investors are risk neutral can be made.

In a world where investors are risk neutral the expected return on all securities is the risk free interest rate. The reason is that risk neutral investors do not require a premium to induce them to take risk. Thus when valuing an option we calculate the expected payoff assuming that the expected return from the underlying asset is the risk free interest rate and use the same risk free interest rate to discount the expected payoff. The price of a European call option can then be written as:

$c = e^{-rT} E[\max(S_T - X, 0)]$ *Equation 2.9*

For details about how this leads to 2.5 please see Hull (1997).

2.3 Implied volatility

In the Black-Scholes framework the option price is a function of six variables; stock price, exercise price, risks free interest rate, time to maturity, dividend rate and volatility. Of these six factors five are observable (at least for an exchange traded stock). The one that is unobservable is the volatility. The volatility that makes the theoretical option price calculated from the Black-Scholes formula equal to the observed market price is called the implied volatility. The implied volatility can easily be found by solving the following equation:

$c_{B\&S} - c_{obs} = 0$ *Equation 2.10*

According to the Black-Scholes model implied volatilities form options should be the same regardless of which options is used to compute the volatility. In practice this is usually not the case. Options on the same underlying asset with different strike prices and maturities yield different implied volatilities. The differences in implied volatility for equity options are sometimes referred to as the volatility skew because the implied volatility typically decreases as the strike price increases. This means that out of the money puts and in the money calls have greater implied volatility than out of the money calls and in the money puts of equivalent maturity. The existence of skewness in the implied volatility is clearly inconsistent with the Black-Scholes model and it means that options are not priced as though the underlying asset follows a geometric Brownian motion and that the underlying asset price is log normally distributed. In the first part of this study I will analyze the short-term relation between implied volatility and future return on the underlying asset.

As I will return to later my hypothesis regarding implied volatility is that it is a contrarian indicator, i.e. when the implied volatility is abnormally high and the market participants are afraid of a downturn the most the implied volatility serves as a contrarian indicator of a short term upturn in the stock market.

The use of implied volatility as a contrarian indicator has been widely spread in the financial markets in latter years. The most famous measure of implied volatility is the VIX-index[17] in the United States. The VIX-index measures the implied volatility of options in the S&P 100 index. The S&P 100 index contains the largest 100 companies in the more famous S&P 500 index.

The VIX-index is used as a tool for measuring investor fear. High readings (>50) mark periods with extremely high implied volatility in the S&P 100 options and consequently maximum investor fear. Readings above 50 have occurred three times since 1998 (October 1998, September 2001 and July 2002). All three times were followed by powerful rallies in the S&P 500 index (and all other major indices as well). The VIX-index is today used in a variety of areas and quotes on the VIX-index are available in real time throughout the day.

As I did mention in the introductory chapter this study will focus on the very short term relations between implied volatility and stock return and no analyze will therefore be made of the longer term relations.

2.4 Implied risk neutral density functions

The implied volatility is not the only information that can be extracted from option prices. Since an option's payoff depends on the future development in the underlying asset the prices of different options contracts reflect the market participants' view of the likelihood that the contract will yield a positive payoff. Thus by studying prices of options on a particular asset with different strike prices but with the same time to maturity they may tell us something about the probability that the market attaches to the asset being within a range of possible prices at some future date.

A popular way of gaining this information from option prices is by estimating the so called implied risk neutral density function. Under given assumptions this function can be interpreted as the market's aggregate probability distribution for the price of the underlying asset at maturity. It may therefore contain information about market expectations e.g. the implied risk neutral density function may tell us whether market participants believe that extreme upward and downward movements are likely to occur.

In the second part of this study I will evaluate whether the properties of the implied risk neutral density functions extracted from OMX-options are related to the upcoming short term return in the OMX-index itself. To do this we need to study the relation between various information parameters in the implied density function and the return on the OMX-index. The parameters of the risk neutral density function that I am going to use in this study are the following:

- The mean (μ)
The expected risk-neutral future value of the OMX-index is equal to the mean of the implied distribution. The mean is sometimes referred to as the first moment of the distribution. When

[17] For more about the VIX-index please see http://www.stricknet.com/vix.

there are no arbitrage opportunities the mean equals the futures price.[18] The mean of the risk-neutral density function *f(x)* is calculated as:

$$\mu = \int_{-\infty}^{\infty} xf(x)dx \qquad \textit{Equation 2.11}$$

A mean above the observed price in the OMX-index i.e. a futures price above the observed OMX-index implies that the markets expectations about future returns are positive over the time horizon of the option contracts.

- The standard deviation (σ)
The second moment of the distribution is the variance. The square root of the variance is the standard deviation. Standard deviation is a measure of the dispersion of the distribution around the mean and is calculated as:[19]

$$\sigma = \sqrt{\int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx} \qquad \textit{Equation 2.12}$$

A high standard deviation for the implied distribution implies that there is high uncertainty about the outcome of the OMX-index at the maturity date of the option contracts. The standard deviation of the implied distribution is related to the implied volatility calculated above, however the exact relationship is complex and I will not go into it further here. The two measures will be treated as separate variables in this study.

- The skewness (Sk)
The third moment of the distribution is the skewness. Skewness characterizes the asymmetry of the implied density function and is calculated as:

$$Sk = \frac{1}{\sigma^3} \int_{-\infty}^{\infty} (x-\mu)^3 f(x)dx \qquad \textit{Equation 2.13}$$

A normal distribution has zero skewness. In a positively skewed distribution the median is lower than the mean i.e. the right tail is in this sense heavier than the left tail. This may suggest that the market is positive about future returns. However, it is important to keep in mind that such positive expectations naturally lead to an upward revision of the mean/futures price and thus consequently the stock price. Because of this, in a positively skewed distribution there is less probability attached to outcomes higher than the mean than to outcomes below the mean.

- Kurtosis (K)
The fourth moment of the distribution is kurtosis. Kurtosis is a measure of how peaked a distribution is and it is calculated as:

---

[18] More about this in the theory chapter of this thesis.
[19] Stig Arild Syrdal, A study of Implied Risk Neutral Density Functions in the Norwegian Option Market, Working paper, Norges Bank (2002) p39.

$$K = \frac{1}{\sigma^4} \int_{-\infty}^{\infty} (x-\mu)^4 f(x) dx \qquad \textit{Equation 2.14}$$

The kurtosis of a normal distribution is equal to three. Excess kurtosis is therefore defined as kurtosis above three. Kurtosis above three implies a larger probability for extreme outcomes compared to the normal distribution i.e. the tails of the distribution is thicker. When the implied density function is calculated there is a lot of uncertainty about the shape of the tails in the distribution. Due to this uncertainty it is not wise to pay too much attention to the kurtosis measure. It will therefore in this study be interpreted with care.

2.5 Earlier research on implied volatility and implied density functions

Since Mark Rubinstein first launched the idea of crashophobia in the late 1980's research on implied volatility and implied density functions have increased a lot. Below I present the conclusions of some of the research in recent years.

Navatte and Villa (2000)[20] found that the implied first moments contain substantial amount of information for future moments of CAC 40[21] returns although this amount decreases with respect to the moment's order. Navatte and Villa also found that the different shapes of the volatility smile are consistent with different distribution of the underlying returns.

Andersen and Andreasen (2000)[22] add Poisson jumps to the standard DVF[23] diffusion models of stock price evolution and compare the evolution of the volatility smile between the DVF model and the jump-diffusion model. They found that in the DVF model the evolution of the volatility smile is highly non-stationary and often counterintuitive. The jump-diffusion model on the other hand produces almost perfect stationary S&P 500 volatility skews. By stationary it is meant that the volatility smile is constant and do not vary with time.

Rosenberg (2000)[24] implements a generalization of implied volatility function models; the dynamic implied volatility function model (DIVF). The DIVF-model separately models the time invariant implied volatility function and stochastic state variables. The model is compared with volatility by strike, volatility by moneyness, time varying volatility by strike and time varying volatility by moneyness models. Using S&P 500 futures options data over the 1988-1997 period the volatility by strike and volatility by moneyness models are found to provide a poor characterization of implied volatilities and option prices. The time varying versions are more accurate but they substantially underpredict implied volatilities for in the money put options. The improvement in model precision for the time varying models indicates that a volatility function defined over exercise price or moneyness changes over time and that current data are useful in predicting the future volatility function.

---

[20] Patrick Navatte and Christophe Villa, The information content of implied volatility, skewness and kurtosis: empirical evidence from long-term CAC 40 options, European Financial Management (2000) Vol. 6 p 41-56.
[21] CAC 40 is the main stock index on the Paris stock exchange.
[22] Leif Andersen and Jesper Andreasen, Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing, Review of Derivatives research (2000), p 231-262.
[23] Deterministic volatility function.
[24] Rosenberg, Implied volatility functions: A reprise, The Journal of Derivatives (2000).

The DIVF model by contrast explains option price behavior in an internally consistent and parsimonious manner by modeling time variation in the implied volatilities as a function of at the money implied volatility, which itself depends on the underlying asset returns. Empirical results indicate that the DIVF model offers substantial improvements in pricing performance compared to implied volatility function models and some improvement compared to time varying implied volatility function models.

Corrado and Su (1997)[25] tested an expanded version of the Black-Scholes option pricing model that accounts for skewness and kurtosis deviations from lognormality in stock price distributions. They found significant negative skewness and positive excess kurtosis in the option implied distribution of S&P 500 index prices. The observed negative skewness and positive excess kurtosis induces a volatility smile when the Black-Scholes formula is used to calculate option implied volatilities across a range of strike prices. By adding skewness and kurtosis adjustment terms the volatility smile is effectively flattened. They conclude that skewness and kurtosis adjustments terms added to the Black-Scholes formula significantly improve accuracy and consistency for pricing deep in the money and deep out of the money options.

Jackwerth and Rubinstein (1996)[26] derive underlying asset risk neutral probability distributions of European options on the S&P 500 index. They found that since the crash of 1987 the risk neutral probability of a three (four) standard deviation decline in the index (about -36 % (-46 %) over a year) is about 10 (100) times more likely than under the assumption of log normality

Nandi (1999)[27] uses a different and very interesting approach. He develops a model of asymmetric information in which an investor has information regarding the future volatility of the price process of an asset and trades an option on the asset. The model relates the level and curvature of the smile in implied volatilities as well as miss-pricing by the Black-Scholes model to net options order flows (to the market maker). It is found that an increase in net options order flows (to the market maker) increases the level of implied volatilities and results in greater miss-pricing by the Black-Scholes model, besides impacting the curvature of the smile. The liquidity of the option market is found to be decreasing in the amount of uncertainty about future volatility that is consistent with existing evidence.

Gemmill and Saflekos (2000)[28] examine whether implied distributions are informative with respect to subsequent stock market moves and to what extent they may be used to reveal investor sentiment. They apply the mixture of two lognormals technique to London's FTSE-100 index options and examine the in sample and out of sample performance of this model in a variety of periods. They found no evidence that the option market anticipate crashes. Neither before the large crash of 1987 nor before the smaller crashes of 1989 and 1997 did the options market become more left skewed. The conclusion is that the index options market reacts to crucial events such as stock market crashes but it does not predict them.

[25] Charles J. Corrado and Tie Su, Implied volatility skews and stock index skewness and kurtosis implied by S&P 500 index option prices The Journal of Derivatives (1997).
[26] Jens Carsten Jackwerth and Mark Rubinstein, Recovering probability distributions from optin prices, The Journal of Finance (1996) No 5 p 1611 ff.
[27] Saikat Nandi, Asymmetric information about volatility: How does it affect implied volatility, option prices and market liquidity?, Review of Derivatives Research (1999).
[28] Gordon Gemmill and Apostolos Saflekos, How useful are implied distributions? Evidence from stock index options, The Journal of Derivatives (2000).

Syrdal (2002)[29] analyzes the potential value of information contained in prices of options on the OBX index at Oslo Stock Exchange. The study shows that there is a high level of uncertainty surrounding the implied density functions extracted from OBX options. The study concludes that using information obtained in OBX-option prices in forecasting future market prices seem to be worthless. Some information about future volatility may be obtained, but not about the direction of future outcomes.

[29] Stig Arild Syrdal, A study of Implied Risk Neutral Density Functions in the Norwegian Option Market, Working paper, Norges Bank (2002) p3.

3 Methodology

To analyze the relation between OMX-options and the return in the stock market I will perform an OLS-regression. To find the explanatory variables in the regression I need to extract the implied volatility and the various moments of the implied density function from the option prices.

3.1 Data collection

Since calculating the implied density function is very time consuming I will restrict this study to the September contracts of 2002. Implied volatility is somewhat faster to calculate due to the help of VBA and this study can therefore be conducted on more strike months and a longer period; August through November. Since I did collect the data separately for the two parts of this study, the data collection will be described separately below.

3.1.1 Implied volatility

This study will be conducted on the OMX-index at the Stockholm Stock Exchange. I will study a period of 86 trading days starting on August the 2nd 2002 and ending on November 29th 2002. All market data for the OMX-index and the included OMX-options are coming from the Stockholm Stock Exchange itself.[30] To be able to study the strike structure of the implied volatility I use data for eight strike prices each day and for call options only. I expect symmetry on average between call and put options, and because calculating implied volatility is time consuming I will exclude put options. I assume symmetry between call and put options due to the put-call parity relationship. The put-call parity shows that the value of a European call with a certain exercise price and exercise date can be deduced from the value of a European put with the same exercise price and exercise date and vice versa.[31] If put-call parity does not hold there are arbitrage opportunities.

I use the strike series that are closest centered around the closing price of the OMX-index i.e. I use four in the money options and four out of the money options. I will further collect data for contracts of two different maturities; contracts that are closest and second closest to expiring.

3.1.2 Implied density function

The study with the implied density function will be conducted on OMX-options that expire in December of 2002. These options were first traded on September 24th 2002. I will therefore study the implied density function for the options contracts that expire in December 2002

[30] Anita Grimstad, OMHEX, HEX Integrated markets/Stockholmsborsen.
[31] John C. Hull, Options, Futures and Other Derivatives (2003) p174.

from September 24th through December 23rd. To be able to get an implied density function that is as accurate as possible I will use all available strike prices of both call and put options when I calculate the density function.

3.1.3 The OMX-index

The OMX-index on the Stockholm Stock Exchange consists of the 30 most traded stock on the exchange. The index is revised every six months when changes in the composition of the index also take place. The OMX-options are standardized options contracts with cash settlement. The options are European i.e. they can only be exercised at maturity.

3.1.4 Further discussion of the data

A major weakness of the OMX-options that are traded on the Stockholm Stock Exchange is that many contracts exhibit low liquidity. Because of the low liquidity relatively large bid-ask spreads are often observed in the data and many option series also exhibit non-synchronous trading. Due to both the bid-ask spread and the non-synchronous trading the options exhibits excess volatility and negative autocorrelation.[32] It may therefore not be correct to use the closing price of the day for the options in the study. Instead I will be using the average of the bid and ask price at the end of the day as a proxy for closing price. This methodology has been used in many prior studies e.g. Syrdal (2002).

Since I only will use OMX-contracts that expire in August through March I do not need to adjust for dividends since stocks do not pay dividends in these months. According to Hull (1997) it is possible to use either calendar days or trading days as a measure for time, in this study I will use calendar days.

As a proxy for the risk free interest rate I will use Swedish 90-day T-bills. Since an investment in a T-bill is only risk free when time to maturity corresponds exactly to the investment horizon and many of the contracts that are studied have time to maturity less than 90 days it would be more correct to interpolate between the 90-day rate and a shorter T-bill rate to obtain the risk free rate. Since the difference between T-bills in this spectrum is limited and the end-effect of interest rates on option prices is very small I will not perform any interpolation but instead simply use the 90-day rate as a proxy for the risk free rate.

3.2 My hypotheses

Below are the hypotheses that I will verify or reject during this study. They are all based on the mean-reversion phenomenon described in the Problem discussion.

1. High and increasing implied volatility predicts a positive short term return

2. A futures price that is below the spot price predicts a positive short term return

[32] Campbell, Lo, McKinlay, The Econometrics of Financial Markets (1997) p101.

3. High standard deviation in the implied density function predicts a positive short term return.

4. A positive skewness in the implied density function predicts a positive short term return.

3.3 Calculating implied volatility

The five of the six inputs in the Black-Scholes formula that are described in the data-section above are used as input to calculate the implied volatility. The sixth component is the volatility itself that is non-observable. However, the price of the option is observable and this will be used to extract the implied volatility from the Black-Scholes formula according to:

$c_{B\&S} - c_{obs} = 0$

where $c_{B\&S}$ is the option price from the Black-Scholes formula and c_{obs} is the observed option price. Simply by varying the volatility in the Black-Scholes formula until the left hand side of the above equation equals zero we will get the implied volatility of the option.

The above calculation can be performed in Excel with the goal-seek function by varying the cell that contains volatility until the above equation is equal to zero. However this is a very time consuming procedure since my analysis contains close to twenty option contracts in each day. Instead I will use the Visual Basic Editor (VBA) in Excel to perform the calculations.

This can be done in several ways. One way is by first creating three functions in VBA; one for the call price, one for the variable d_1 on the call price function and one function for the variable d_2 in the call price function.

The above goal-seek function can then be replicated by creating a user defined function called BSOptionISDGGoalSeekNR. This Goal-seek function ensures that each new guess of the volatility takes account of the current error (the distance between the current price using the guessed volatility and the observed price) and also the slope of the option price with respect to volatility. Using the slope to improve the accuracy of subsequent guesses is known as the Newton-Raphson method.[33] It can easily be programmed within a loop until the observed option price is reached within a certain specified tolerance. The functions are more specifically described in Jackson & Staunton (2003).

Another approach is to use a VBA-program that repeatedly applies the above solver-function. This can be done by simply program the solver function in VBA within a loop that continues until all option series are solved.[34] More about this procedure can be found in Appendix A. This approach was the one that I finally decided to use in this study but I also tried the Newton-Raphson method described above and it works equally well.

During this thesis I discovered many other ways to calculate the implied volatility in VBA that work equally well. For those interested I can recommend Jackson & Staunton (2003)[35],

[33] Mary Jackson and Mike Staunton, Advanced modelling in finance using Excel & VBA (2001) p 211.
[34] Simon Benninga, Financial Modelling (2000) p 206.
[35] Mary Jackson and Mike Staunton, Advanced modelling in finance using Excel & VBA (2001) p 217 ff.

Benninga (2000)[36] or Sengupta (2004)[37]. They all describe various methods to model options in VBA.

3.4 Calculating the Implied risk neutral density function

The second part of this study will be to study how the information in the implied risk neutral density function is related to upcoming stock return. To do this I will study the OMX-options contracts that expire in December of 2002. These contracts did start to trade at the Stockholm exchange on September 24th. I will therefore study the contracts in the period September 24th to the expiration day December 23rd. There are three main ways to calculate the implied distribution; the single lognormal model (SLN), the double lognormal model (DLN) and the Smoothed implied volatility smile method (SPLINE). In this study I will use the double lognormal model exclusively, however understanding the double lognormal model is easier with knowledge about the single lognormal model. I therefore start with explaining the single lognormal model. Most of this chapter is based on Syrdal (2002) so if you are interested in more detailed information I recommend that paper.

3.4.1 The Single Lognormal model (SLN)

In chapter 2 the Black-Scholes model for valuing options is reviewed. This model assumes that the price of the underlying asset is lognormally distributed and that the return of the asset is normally distributed with constant variance. The lognormal distribution for a stochastic variable x can be described by two parameters, α and β as:[38]

$$L(x \mid \alpha,\beta) = \frac{1}{x\beta\sqrt{2\pi}} e^{-(\ln x-\alpha)^2/(2\beta^2)} \qquad \textit{Equation 3.1}$$

In the Black-Scholes model options are priced as if investors are risk neutral by setting the expected rate of return on the underlying asset, μ equal to the risk free interest rate, r. The parameters of the risk-neutral lognormal distribution for the underlying asset at maturity can then be expressed as:

$$\alpha = \ln S_0 + (r - \frac{\sigma^2}{2})T \qquad \textit{Equation 3.2}$$

$$\beta = \sigma\sqrt{T} \qquad \textit{Equation 3.3}$$

where S_0 is the current price of the underlying asset, T is the remaining time to expiration and σ is the volatility of the underlying asset.

[36] Simon Benninga, Financial Modelling (2000) p 303 ff.
[37] Chandan Sengupta, Financial modelling using Excel and VBA (2004) p 583 ff.
[38] Syrdal, A study of implied risk neutral density functions…, Working paper Norges Bank (2002) p 27.

Since the expected return on the underlying asset is equal to the risk-free interest rate, the expected future value of the underlying asset at maturity must equal $S_0 e^{rT}$. The mean of the lognormal distribution is given by $e^{\alpha+\beta^2/2}$. Thus the current price of the underlying asset S_0 can be found by:

$$S_0 e^{rT} = E\left[L(S_T|\alpha,\beta)\right] = e^{\alpha+\frac{1}{2}\beta^2} \qquad \textit{Equation 3.4}$$

$$\rightarrow S_0 = e^{-rT} e^{\alpha+\frac{1}{2}\beta^2} \qquad \textit{Equation 3.5}$$

If we then substitute the expression for S_0 and the expression for α and β given in Equation 3.2 and 3.3 into the Black-Scholes formulas given in equation 2.6 and 2.7 the price of a European call and put option with strike price X can be written as:

$$c(X,\tau) = e^{-rT}\left[e^{\alpha+\frac{1}{2}\beta^2} N(d_1) - XN(d_2)\right] \qquad \textit{Equation 3.6}$$

$$p(X,\tau) = e^{-rT}\left[-e^{\alpha+\frac{1}{2}\beta^2} N(-d_1) + XN(-d_2)\right] \qquad \textit{Equation 3.7}$$

where

$$d_1 = \frac{-\ln(X) + \alpha + \beta^2}{\beta} \qquad d_2 = d_1 - \beta$$

Equations 3.6 and 3.7 are the general Black-Scholes formulas expressed in terms of the parameters of the underlying lognormal distribution at maturity, α and β. Since I have assumed a proxy for the risk-free interest rate (3-month T-Bill) the only unknown parameters in the above expressions are α and β. These parameters can be estimated by minimizing the squared deviation between the observed prices and the theoretical option price calculated from 3.6 and 3.7. This is explained more in detail below, in the section about the double lognormal model.

3.4.2 The Double Lognormal model (DLN)

Because options generally are not priced as though the price of the underlying asset is lognormally distributed we need a more flexible density function than the SLN. A widely used method is to assume that the distribution for the underlying asset is a weighted sum of two lognormal distributions. For options that are available in a large number of strike prices it is possible to weight more than two lognormal density functions and thereby get an even more accurate distribution. However in a market like the Swedish options market with a rather limited number of strike prices available it is only possible to weight two distributions.

I will therefore assume that the underlying distribution is a combination of only two lognormal density functions. The double lognormal distribution is described by five parameters: two parameters for each lognormal distribution (α_1,β_1 and α_2,β_2), and a weighting parameter (θ) which describes the relative weight of each distribution. The risk neutral lognormal distribution can be written as:

$q(S_T) = \theta * L(S_T \mid \alpha_1,\beta_1) + (1- \theta) * L(S_T \mid \alpha_2,\beta_2)$ *Equation 3.8*

In the section about the Single lognormal model I concluded that the distributional parameters (α and β) can be estimated by minimizing the squared deviation between the observed option prices and the theoretical prices obtained from equations 3.6 and 3.7 i.e. the theoretical option prices from the risk-neutral valuation. The expressions for the call and put price in the Double lognormal model are:

$$c(X,T) = e^{-rT}\left\{\theta\left[e^{\alpha_1+\frac{1}{2}\beta^2{}_1}N(d_1) - XN(d_2)\right] + (1-\theta)\left[e^{\alpha_2+\frac{1}{2}\beta^2{}_2}N(d_3) - XN(d_4)\right]\right\}$$

Equation 3.9

p(X,T)=

$$e^{-rT}\left\{\theta\left[-e\alpha_1 + \frac{1}{2}\beta^2{}_1 N(-d_1) + XN(-d_2)\right] + (1-\theta)\left[-e\alpha_2 + \frac{1}{2}\beta^2{}_2 N(-d_3) + XN(-d_4)\right]\right\}$$

Equation 3.10

where

$$d_1 = \frac{-\ln(X) + \alpha_1 + \beta_1^{\,2}}{\beta_1} \qquad d_2 = d_1 - \beta_1$$

$$d_3 = \frac{-\ln(X) + \alpha_2 + \beta_1^{\,2}}{\beta_2} \qquad d_4 = d_3 - \beta_2$$

If we compare the option prices in 3.9 and 3.10 with the expressions obtained for the single lognormal model (3.6 and 3.7) we see that under the double lognormal assumption the option price is a weighted sum of two Black-Scholes solutions.

In absence of arbitrage opportunities the currently observed futures price (F) for the underlying asset should equal the mean of the risk neutral density function. The mean of the double lognormal distribution is a weighted sum of the individual means. Hence, the following relationship must be satisfied:

$$F = \theta e^{\alpha_1 + \frac{1}{2}\beta_1^2} + (1-\theta)e^{\alpha_2 + \frac{1}{2}\beta_2^2}$$ *Equation 3.11*

One way of implementing 3.11 is to impose it as a constraint in the minimizing problem. However in this study I will use a more general method suggested by Syrdal, I will add the squared deviation between the futures price and the mean of the distribution to the minimizing problem.

As described earlier I will use the 90-day Swedish T-bill as a proxy for the risk free interest rate. We then only have five distributional parameters (α_1, α_2, β_1, β_2 and θ) that need to be estimated in order to obtain the implied risk-neutral density function for the Double lognormal model. These parameters are recovered by minimizing the squared deviation between the observed option prices for a given maturity, together with the squared mean-futures price deviation. The minimizing problem can be written as:

$$\min\left\{\sum_{i=1}^{m}(c_i - c_i^*)^2 + \sum_{j=1}^{n}(p_j - p_j^*)^2 + \left[\theta e^{\alpha_1 + \frac{1}{2}\beta_1^2} + (1-\theta)e^{\alpha_2 + \frac{1}{2}\beta_2^2} - F\right]^2\right\}$$

Equation 3.12

where

c_i^*, p_j^* = observed option prices
c_i, p_j = theoretical option prices estimated from 3.9 and 3.10
m = number of call options in the data set
n = number of put options in the data set

In practice this minimizing problem is simplified. Bahra (1997) shows that if the price of the underlying asset is assumed to follow a lognormal distribution the price of a European call and put option can be expressed analytically.[39] Using the analytical price expression instead of the risk-neutral valuation equations will simplify the minimizing problem a great deal.

[39] Stig Arild Syrdal, A study of Implied Risk Neutral Density Functions in the Norwegian Option Market, Working paper, Norges Bank (2002) p29.

A weakness that the double lognormal model has is that it sometimes produces a density function which is characterized by a sharp spike. The reason is that one of the two lognormal distributions is estimated to have a very small standard deviation. It also happens that the optimization procedure fails in finding any solution for a particular dataset. To overcome these difficulties I have put the following restrictions, suggested by Syrdal on the β values:

$$0{,}15 < \frac{\beta_1}{\beta_2} < 4$$

The minimization problem is being solved with the problem solver in Excel. The target cell is the cell containing the sum of the squared deviations between observed option prices and theoretical option prices according to 3.12. The cells that are used to minimize 3.12 are the cells containing the values of α_1, α_2, β_1, β_2 and θ. The restrictions I impose on the solution are in addition to the one regarding the quote of the Betas that θ (the weight) should be between zero and one.

3.5 Regression analysis

This study is focusing on the short term relations between implied volatility and return on the one hand and on the other hand the short term relations between the implied density function and return. This is an important difference to many other studies that mainly study the changes in implied volatility or implied density functions due to large events. This study will focus on the day to day changes in implied volatility and the implied density function and I will analyze how these changes are related to future changes in return of the OMX-index. I will use Eviews in all regressions and heteroskedasticity and autocorrelation will be tested for and if necessary standard errors corrected in all regressions.

Since the sample size for particularly the regression with the implied density function is somewhat small it is important to only adjust for heteroskedasticty if it is present. In small sample sizes robust t-statistics can have distributions that are not very close to the t-distribution and that could throw off the inference.[40]

I will run two different regressions; one with implied volatility as explanatory variable and one with the various moments of the implied density function as explanatory variables.

3.5.1 Controlling explanatory variables

Besides the explanatory variables that are of interest for the study I will add two more explanatory variables for controlling purposes to all of the regressions in the study; a decreasing time trend and the return in day *t*.

The time trend will be included because the volatility of the implied density function is a decreasing function of maturity. The implied density function measures the probability of different outcomes for the OMX-index at the strike day. The uncertainty about the outcome at the strike day must intuitively decrease as the contract approach the strike day, i.e. the

[40] Jeffrey M. Wooldridge, Introductory Econometrics (2003) p261.

volatility of the implied density function must decrease as well. To be able to compare the volatility of the implied density over time this falling trend must be accounted for. I will do this by simply adding a negative time trend that decreases with one for each day as we approach maturity.

The return of the OMX-index in day *t* is the second controlling variable. The logic behind including this variable is that the various moments of the implied density function in day *t* might be correlated with the return in the same day, day *t*. To make sure that we only catch the lagged effect of the implied density function on return i.e. the effect on return in day *t+1* I will include the return in day *t* as an explanatory variable.

3.5.2 Implied volatility as explanatory variable

In this regression focus will be on the day to day change in implied volatility i.e. the relative change in implied volatility and the absolute value of the implied volatility. For this purpose I will use the call contract that are closest to at the money and have the shortest time to maturity. For example if the OMX-index is closing at 518 on September 15th I will use the call contract with a strike price of 520 that expires on the fourth Friday of September. The implied volatility for each day for the specified contract according to the above restriction will then be one of the explanatory variables.

The other explanatory variable will be the relative change in implied volatility from day to day according to:

$$\Delta\sigma = \ln\frac{\sigma_t}{\sigma_{t-1}}$$

where σ represents implied volatility and $\Delta\sigma$ represents the change in implied volatility from day *t-1* to day *t*.

I will run two different regressions with these explanatory variables; one with the return in day *t+1* of the OMX-index as dependent variable alone and one with the accumulated return in day *t+1*, *t+2* and *t+3* of the OMX-index as dependent variable. This way I will obtain an indication of what prediction power implied volatility in day *t* is having on the return of the OMX-index in the next day and the three upcoming days together. The two regressions are specified according to:

$$r_{t+1} = \alpha + \beta_1 R_t + \beta_2 \sigma_t + \beta_3 \Delta\sigma + \beta_4 TT + e$$

$$r_{t+1,2,3} = \alpha + \beta_1 R_t + \beta_2 \sigma_t + \beta_3 \Delta\sigma + \beta_4 TT + e$$

where r_{t+1} is the return in OMX-index in day *t+1* and $r_{t+1,2,3}$ is the accumulated return in the OMX-index for days *t+1*, *t+2* and *t+3*. $\Delta\sigma$ is the change in implied volatility from day *t-1* to day *t* and *TT* is a time trend. I will use Eviews for all regressions and I will test for heteroskedasticity in the relations.

β_1 and β_2 will be tested with a t-test to investigate the hypothesis that there is a significant prediction power on the upcoming return from implied volatility. The hypotheses are the following:

H_0: β_1 and/or $\beta_2 = 0$
H_1: β_1 and/or $\beta_2 \neq 0$

If the null hypothesis is accepted implied volatility does not predict future return and if the null hypothesis is rejected implied volatility does predict future return.

3.5.3 Moments of the implied density function as explanatory variables

In these regressions I will include the second, third and fourth moment of the implied density function i.e. volatility, skewness and kurtosis. Since I do not run these regressions for the whole period as I run the above regression I will also here include the variables for implied volatility and relative change in implied volatility. The contracts that are used to calculate the implied density function are the contracts that expire on December 23rd. They were first traded on September 24th and the period between these dates are used in the regression. This is a somewhat shorter period than above.

Just as with the implied volatility above I will run two regressions; one with the return in day *t+1* as dependent variable and one with the accumulated return for day *t+1*, *t+2* and *t+3* as dependent variable. The two regressions are then specified according to:

$$r_{t+1} = \alpha + \beta_1 R_t + \beta_2 F_t + \beta_3 \sigma_t + \beta_4 \Delta\sigma + \beta_5 Vol_t + \beta_6 Sk_t + \beta_7 Kur_t + \beta_8 TT + e$$

$$r_{t+1,2,3} = \alpha + \beta_1 R_t + \beta_2 F_t + \beta_3 \sigma_t + \beta_4 \Delta\sigma + \beta_5 Vol_t + \beta_6 Sk_t + \beta_7 Kur_t + \beta_8 TT + e$$

where r_{t+1}, $r_{t+1,2,3}$, σ, $\sigma\Delta$ are defined as above. Vol_t, Sk_t *and* Kur_t denotes the volatility, skewness and kurtosis respectively of the implied density function. The β-values will be tested as above to determine if they are significant separated from zero.

A positive β_1 indicates that the implied volatility in day *t* predicts a positive return in day *t+1/t+1,2,3*. A positive β_2 indicates that an increase in implied volatility from day *t-1* to day *t* predicts a positive return in day *t+1/t+1,2,3*.

4 Empirical results

In this chapter the results from the regressions are presented. I will start with the result for the regressions with implied volatility as main explanatory variable in section 4.1 and continue with the regressions that adds the moments of the implied density functions as explanatory variables in section 4.2.

4.1 Relation between implied volatility and short term return

Table 4.1 shows the result of the regression with return in day *t+1* as dependent variable. The regression was tested for heteroskedasticity and there were no sign of any heteroskedasticity present. The Durbin-Watson value is close to 2 i.e. there are no sign of first order autocorrelation.

Table 4.1: Regression results with return on OMX-index, day *t+1*, as dependent variable.

Variable	Coefficient	t-stat	p-value
Return, day *t*	-0.014	-0.121	0.904
Implied Volatility	0.072	1.871	0.065
Δ Implied Volatility	-0.002	-0.100	0.921
Timetrend	0.000	-1.704	0.092
Constant	-0.058	-1.569	0.122
R-squared	0.0641		
Durbin-Watson	1.9804		
F-statistic	1.369		
P-value	0.252		

The regression shows that implied volatility in day *t* is positively related to the return of the OMX-index in day *t+1*. The coefficient is significant at the 6.5 % level. Implied volatility is measured in "level" value and return in "log" value which means that a higher implied volatility by 0.1 predicts a higher OMX-index return in the following day by 0.72 % under the assumption that the remaining explanatory variables are unchanged; a rather strong one day impact.

The relative change in implied volatility between day *t-1* and day *t* on the other hand does not seem to predict the return of the OMX-index in day *t+1*. Both the statistical and economic significance is small. The R-squared of the regression is 6.4 % which at first sight might seem low but on the other hand we have to remember that there are rather few variables in the regression and we are only measuring the impact on one day.

My hypothesis about implied volatility is economically verified. A high implied volatility in day *t* means that options are expensive i.e. the market participants are expecting a large move in either direction and therefore require a higher premium to write options. According to the problem discussion of this thesis high implied volatility is more associated with nervousness for large downturns than euphoria for large upturns. This can be interpreted as when the implied volatility is high and options are expensive the majority of the market participants believe the market will fall. When this belief is widely spread (shown in form of high implied volatility) all negative news are already accounted for and the market instead turns up in the short term. Implied volatility therefore serves as a contrarian indicator of the short term return.

The use of implied volatility as a contrarian indicator has been widely spread in the financial markets in the last years. In the United States the so called VIX-index[41] is often used to forecast the S&P 500 index. The results of the regression on the OMX-index option implied volatilities are therefore in accordance with the evidence from the VIX-index as a contrarian indicator.

Table 4.2 shows the results for the regression with the accumulated return on OMX-index for day *t+1*, *t+2* and *t+3*. White's test did show no sign of heteroskedasticity but the Durbin-Watson values indicate that there is a strong presence of positive autocorrelation. I therefore corrected the regression with the Newey-West method in Eviews. The values in the table below are adjusted with the Newey-West method.

Table 4.2: Regression results with accumulated return on OMX-index, over days *t+1*, *t+2* and *t+3*, as dependent variable.

Variable	Coefficient	t-stat	p-value
Return, day *t*	-0.095	-0.588	0.558
Implied Volatility	0.131	1.853	0.068
Δ Implied Volatility	0.028	0.907	0.367
Timetrend	0.000	-2.363	0.021
Constant	-0.031	-1.170	0.246
R-squared	0.136		
Durbin-Watson	0.643		
F-statistic	3.068		
P-value	0.021		

When the returns for the three days following day *t* are accumulated, the explanatory power of the regression increases; the R-squared is now close to 14 %. The impact from implied volatility almost doubles while the significance remains the same. A coefficient of 0.13 implies that when the implied volatility is higher by 0.1 this increases the return in the following three days accumulated by 1.3 %. The affect from the relative change in implied volatility also increases and the relation is positive in this regression, however the relation is not significant. The overall significance of the regression also increases and the p-value is 0.021, which is a rather strong significance and it seems as the explanatory variables in the regression explain a significant part of the return in the OMX-index in the upcoming days.

[41] More information about the VIX-index can be found in the theory chapter.

4.2 Relation between the implied density function and short term return

Table 4.3 shows the results for the regression with the return of OMX-index in day *t+1* as dependent variable. When I did White's test for heteroskedasticity there were no signs of heteroskedasticity. There is some tendency for negative autocorrelation but the Durbin-Watson values are within a reasonable range for accepting the hypothesis of no first order autocorrelation.

Table 4.3: Regression results with return on OMX-index, day *t+1, t+2* and *t+3* as dependent variable.

Variable	Coefficient	t-stat	p-value
Return, day *t*	-0.099	-0.780	0.439
Future	0.002	0.748	0.458
Volatility	0.001	0.993	0.325
Skewness	0.035	2.744	0.008
Kurtosis	0.006	1.813	0.075
Implied Volatility	0.195	2.837	0.006
Δ Implied Volatility	-0.029	-1.015	0.314
Timetrend	-0.002	-2.831	0.006
Constant	-0.058	-1.569	0.122
R-squared	0.407		
Durbin-Watson	2.182		
F-statistic	4.813		
P-value	0.000		

The explanatory power of the first through the fourth moment of the implied density function together with the implied volatility is very high. The R-squared is close to 40 % which means that 40 % of the return in OMX-index in day *t+1* can be explained purely by the information from option prices in day *t*. Looking at the individual figures more closely we see that the effect from the futures price in day *t* is rather small and insignificant. The tendency is though that when the futures price is above the index price in day *t* this predicts a positive return in the OMX-index in the next day. This is in conflict with my hypothesis regarding futures but the result is not significant, either economically or statistically.

The effect from volatility in the implied density function is rather limited and insignificant. However, it has the same sign as the implied volatility, something that is reasonable to believe.

The skewness in the implied density function is positively related to return in OMX-index in day *t+1*. This is in accordance with my hypothesis that in a positively skewed distribution there is less probability attached to outcomes higher than the mean than to outcomes below the mean due to the fact that positive expectations naturally lead to an upward revision of the stock price. This means that when the distribution is positively skewed market participants believe that the market will fall, and when the large majority of the participants believe this the market goes the other way i.e. skewness is a contrarian indicator. This result is strong with a 1.7 % significance level.

Kurtosis is also positively related to the return in day *t+1*. As discussed earlier the reliability of kurtosis as a measure of future return is poor and the variable should therefore be interpreted with care. However, the regression shows that excess kurtosis (kurtosis >3) is positively related to the return in day *t+1*. This means that an implied density function with fatter tails than the normal distribution is positively related to the return in the following day.

The two variables for implied volatility show the same result as in the regression with the two of them exclusively. Since this regression is not ran on exactly the same time period I thought it was interesting to include them here as well and since the variable for the absolute value of the implied volatility is significant also in this regression this strengthens my results from regression (1).

Table 4.4 shows the results for the regression with the accumulated return in day *t+1*, *t+2* and *t+3* as dependent variable. Neither in this regression did we find any sign of heteroskedasticity but the Durbin-Watson values show that there is strong presence of positive first order autocorrelation between the error terms. This is expected since we are adding three days of return. I correct for autocorrelation with the Newey-West method in Eviews and the table below shows the result after Newey-West correction.

Table 4.4: Regression results with return on OMX-index, day *t+1, t+2* and *t+3* as dependent variable.

Variable	**Coefficient**	**t-stat**	**p-value**
Return, day *t*	-0.138	-0.854	0.397
Future	0.002	0.949	0.347
Volatility	0.000	-0.300	0.765
Skewness	0.045	1.352	0.182
Kurtosis	0.007	0.716	0.477
Implied Volatility	0.236	1.574	0.121
Δ Implied Volatility	0.009	0.261	0.795
Timetrend	-0.003	-1.490	0.142
Constant	0.037	0.341	0.734
R-squared	0.622		
Durbin-Watson	0.876		
F-statistic	11.128		
P-value	0.000		

Compared to the regression with only the return in day *t+1* as dependent variable the R-squared is much higher here i.e. the implied density function and implied volatility explain a much larger part of the return in the three upcoming days than in the first upcoming day exclusively.

For skewness, kurtosis and implied volatility the affect is also higher on the accumulated three days of return than on the first day itself. However, the effect is no longer significant. The effect from the futures price and the volatility of the implied density function remain pretty much unchanged.

5 Conclusions

I start this chapter by answering the hypotheses I put up in the Methodology chapter. After this follows a discussion of the results and I end with some suggestions for further research.

5.1 Hypotheses

1. High and increasing implied volatility predicts a positive short term return

This study shows that high implied volatility is positively related to return in the upcoming days. This is true both for the first day alone and when accumulating the first three days. However; increasing implied volatility shows no significant effect on the return, neither in the first day or the three first days accumulated.

2. A futures price that is below the spot price predicts a positive short term return

This hypothesis is not verified. Futures prices are not significantly related to upcoming return.

3. High standard deviation in the implied density function predicts a positive short term return.

This hypothesis is not verified. The standard deviation does not significantly affect the upcoming return.

4. A positive skewness in the implied density function affects short term return positively.

This hypothesis is strongly verified in the study over the first day alone. The same tendency is true for the first three days accumulated but this result is not significant.

5.2 Discussion of Results

This study has analyzed the short term relation between information implied by option prices and the upcoming return in the OMX-index. It shows that high implied volatility of at the money options affect the return in the OMX-index positively in the next three days. My logic behind this argument as I stated in the problem discussion therefore seems to be true i.e. high implied volatility occurs at the end of a period of negative return due to extreme nervousness of continuous negative return. Since stock returns are mean-reverting over time the mean reversion turns stock returns positive in the short run and implied volatility falls. It is usually very difficult by simply looking at the stock return to see what is considered to be a long period of negative return. To measure the level of nervousness in the market implied volatility of options seems as an interesting tool since it conveys information about the market participants' belief about future return.

The relation is also strong for the third moment of the implied density function i.e. skewness. A positively skewed distribution i.e. a distribution with less probability attached to outcomes above the mean is positively related to the return in the upcoming three days. This is because of the same logic as for the argument above concerning implied volatility i.e. skewness measures the degree of probability to return above the mean. When this probability is low i.e. positive skewness, stock returns have usually been negative for several days and due to the fact that stock returns will mean revert (in this case up) this positive skewness is related to upcoming positive short term return.

The hypothesis regarding the standard deviation in the implied density function could not be verified in this study. This could be because a distribution with a high standard deviation is not related any specific direction of the upcoming return.

This study shows in many ways much more evidence for a relation between information in option prices and future returns than earlier research. One reason for this could be that I am using a rather short time period in my study, maybe the result would have been different if I would have doe the study on a longer time period. Because of the short time period studied the study lacks some reliability. Another reason for the more strong relations obtained in this study might be that I have only studied short term, day to day relations. Most earlier research like Syrdal (2002) that concludes that "information in OBX-options are worthless in forecasting the market" focus totally on the implied density function's relation to long term return and large market crashes.

A third reason for different results could very well be that the relation is not the same for all stock markets. My study is conducted on the Swedish OMX-index in opposition to Syrdal who did his study in the Norwegian OBX-index. The earlier research that my study seems to coincide with the most is the study performed by Navatte & Villa (2000) on the Paris CAC 40 index. They conclude that there is a relation between the various moments of the implied density function and future stock returns, however the relation is declining with respect to the moment's order.

Finally I think that the results in this study are interesting and provide some ground for future research of short term relations between information of option prices and stock return. The major weakness of this study remains that the analysis is conducted on a too short period of time to be able to call the study truly reliable.

5.3 Suggestions for further research

I sum up with some suggestions for further research that I have come up with during the study.

- Perform the study on a much longer time period, or divide the study in to several time periods and compare the results

- Do the study on stock options instead of OMX-options

- Do the study on weekly returns instead of daily returns

- Analyze the sensitivity of the results to different volatility in stock/index return

- Analyze several markets and compare the results between markets with different correlation or industry structure

6 References

6.1 Articles

Andersen, Leif & Andreasen, Jesper, *Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing*, Review of Derivatives research 2000, No.4, p 231-262

Balvers, Ronald, Wu, Yangru, Gilliland, Erik *Mean reversion across National Stock Markets and Parametric Contrarian Investment Strategies*, The Journal of Finance (2000) Vol. 60 No. 2 p. 745-772

Corrado, Charles J. & Su, Tie *Implied volatility skews and stock index skewness and kurtosis implied by S&P 500 index option prices* The journal of Derivatives (1997)

Gemmill, Gordon & Saflekos, Apostolos *How useful are implied distributions? Evidence from stock index options*, The journal of Derivatives, spring 2000

Jackwerth, Jens Carsten,& Rubinstein, Mark, *Recovering Stochastic Processes from Option Prices*, The Journal of Finance (1996) Vol. 51, No. 5, p1611-1631

Nandi, Saikat, Asymmetric information about volatility: *How does it affect implied volatility, option prices and market liquidity?*, Review of Derivatives research 1999, No.3 p 215-236

Navatte, Patrick & Villa, Christophe, *The information content of implied volatility, skewness and kurtosis: empirical evidence from long-term CAC 40 options*, European financial management 2000 Vol. 6, No.1, p 41-56

Syrdal, Stig Arild, *A study of Implied Risk Neutral Density Functions in the Norwegian Option Market*, Working paper Securities Markets and International Finance Department Norges Bank 2002

6.2 Literature

Benninga, Simon, *Financial Modeling*, Second edition, Massachusetts Institute of Technology 2000

Campbell, John Y., Lo, Andrew W. and MacKinlay, A. Craig, *The Econometrics of Financial Markets*, Princeton University Press, United States of America 1997

Eiteman, David K., Stonehill, Arthur I., Moffett, Michael H., *Multinational Business Finance*, Tenth edition, Pearson education Inc 2004

Hull, John C., *Options, Futures and Other Derivatives*, Pearson education Inc, Upper Saddle River, New Jersey 2003

Jackson, Mary & Staunton, Mike, *Advanced modeling in finance using Excel & VBA*, John Wiley and Sons Ltd., The Atrium, Southern Gate, Chichester, England 2001

Oxelheim, Lars, *Financial Markets in Transition*, Norstedt & Söner 1996

Sengupta, Chandan, *Financial Modeling using Excel and VBA*, John Wiley and Sons Inc., Hoboken, New Jersey 2004

Wooldridge, Jeffrey M., *Introductory Econometrics, A modern approach*, Second edition, Thomson Learning, USA 2003

Appendix

The purpose of this Appendix is to give a short presentation of how the implied volatility and the implied density function have been calculated. The figures shown below are a small abstract of the Excel-sheets that have been used but the methodology is still evident.

Appendix A

Below are an extract of the Excel-sheet that is used to calculate implied volatility in VBA. The VBA formula is also shown below. The column *c*-c* is set to zero by varying the implied volatility (column σ).

DATE	SERIE	X	S	Ränta	T	T/365	Exp	BID	ASK	Average	σ	c*-c	B&S	d1	d2	N(d1)	N(d2)	CLOSE	HIGH	LOW
20911	OMX2I570	570	532,71	0,042	16	0,043835616	-0,002	2	2,75	2,375	0,290 –	0	2,374999	-1,05236874	-1,11315593	0,146	0,133	2	2,75	1,75
20911	OMX2J500	500	532,71	0,042	44	0,120547945	-0,005	45	47,5	46,25	0,347 –	0	46,249999	0,628018775	0,507487319	0,735	0,694	0	0	0
20911	OMX2J510	510	532,71	0,042	44	0,120547945	-0,005	38	40,5	39,25	0,341 –	0	39,25	0,469657051	0,351167762	0,681	0,637	39	39	39
20911	OMX2J520	520	532,71	0,042	44	0,120547945	-0,005	31,5	33,75	32,625	0,333 –	0	32,625	0,310584727	0,195038705	0,622	0,577	0	0	0
20911	OMX2J530	530	532,71	0,042	44	0,120547945	-0,005	25,5	27,75	26,625	0,325 –	0	26,625	0,146504003	0,03371997	0,558	0,513	0	0	0
20911	OMX2J540	540	532,71	0,042	44	0,120547945	-0,005	20,25	22,5	21,375	0,318 –	0	21,375	-0,0218547	-0,1324217	0,491	0,447	22	22	19,5
20911	OMX2J550	550	532,71	0,042	44	0,120547945	-0,005	15,75	17,75	16,75	0,312 –	0	16,75	-0,19437448	-0,3025518	0,423	0,381	16	16	16
20911	OMX2J560	560	532,71	0,042	44	0,120547945	-0,005	10	16	13	0,308 –	0	12,999999	-0,36686722	-0,47369264	0,357	0,318	10,75	10,75	10

SolverOK Setcell:="M2", Maxminval:=3, Valueof:="0", Bychange:="L2"
Solversolve Userfinish:=True

…where *M2* is the cell that represents *c*-c* and *L2* is the cell that represents implied volatility. The formula is then programmed to repeat itself within a loop for all cells in the *M* and *L* columns. The implied volatilities that are obtained in column L are then used in the regressions.

Appendix B

The same inputs as above are then used to calculate the implied density function. The squared deviations between the theoretical option prices and the observed options prices are minimized by varying the parameters below that are described in the Methodology chapter.

α1	6,355064261	**Medel 1**	**Medel2**	**Medel DL**
β1	0,077610284	577,134845	228,19	575,30166
α2	5,429987199			
β2	0,019402571			
θ	0,994746494			
Restric	3,999999922			
Start	6,347073924			
Målcell	24,50066225			

The various moments of the implied density function are then calculated:

S	**LogN1**	**LogN2**	**DoubleLog**	**Volatility**	**Skewness**	**Kurtosis**
250	1,82598E-27	1,23E-06	6,4432E-09			
251	3,15592E-27	4,53E-07	2,3818E-09	0,0002505	-0,0812382	26,34568
252	5,42836E-27	1,61E-07	8,4755E-10	8,859E-05	-0,0286411	9,259711
253	9,29275E-27	5,53E-08	2,9045E-10	3,017E-05	-0,0097242	3,13412
254	1,58334E-26	1,83E-08	9,5899E-11	9,9E-06	-0,0031809	1,022032
255	2,68523E-26	5,81E-09	3,0521E-11	3,131E-06	-0,001003	0,321249
256	4,53298E-26	1,78E-09	9,3678E-12	9,551E-07	-0,000305	0,097373
257	7,61729E-26	5,28E-10	2,7739E-12	2,81E-07	-8,946E-05	0,028474
258	1,27424E-25	1,51E-10	7,9283E-13	7,982E-08	-2,533E-05	0,008037
259	2,12207E-25	4,17E-11	2,1881E-13	2,189E-08	-6,924E-06	0,00219
260	3,51835E-25	1,11E-11	5,8338E-14	5,8E-09	-1,829E-06	0,000577
261	5,80781E-25	2,86E-12	1,5031E-14	1,485E-09	-4,667E-07	0,000147

Volatility	44,778491
Skewness	0,3563611
Kurtosis	3,151515