Modeling implied correlation matrices using option prices

SOFIE EKLUND

RANDA ESTAIFO

Modeling implied correlation matrices using option prices

SOFIE EKLUND
RANDA ESTAIFO

Abstract

In the process of calculating a fair value it is preferable to price the asset from observable market data. Some assets are valued using variables which can not be directly observed in the market but are instead implied from observable market data. One such variable is the correlation between assets. The purpose of this thesis is to model correlations between stocks based on observable market data. Three different approaches are used to construct implied correlation matrices on OMXS30. All matrices are constructed using implied volatilities from the option market. The methods are then compared in order to determine which method that generates the most reliable implied correlation matrix. This is done by looking at deviations from counterparty prices on basket options. The used basket options have two different types of underlying autocallable products; Phoenix Autocall and Autocall Uncapped. It was found that the method with an equicorrelation matrix had the smallest deviations from the counterparty price in a majority of the tested cases. Another result was that the implied correlation matrices performed better on the basket options with Autocall Uncapped than Phoenix Autocall as underlying. An interesting topic for further research is to examine other markets but also to study the methods when more than one market is considered.

Sammanfattning

I processen att skapa en rättvis värdering av en tillgång föredras observerbar marknadsdata. Vissa tillgångar värderas med hjälp av variabler som inte direkt kan observeras på marknaden men som kan skapas implicit från observerbar marknadsdata. En sådan variabel är korrelationen mellan tillgångar. Syftet med detta examensarbete är att modellera korrelationer mellan aktier baserat på observerbar marknadsdata. Tre olika tillvägagångssätt används för att konstruera implicita korrelationsmatriser på OMXS30. Alla matriser är konstruerade med hjälp av implicita volatiliteter från optionsmarknaden. Metoderna jämförs sedan för att bestämma vilken metod som genererar den mest tillförlitliga implicita korrelationsmatrisen. Methoderna jämförs sedan genom att granska avvikelser från motpartspriser på optionskorgar. De använda optionskorgarna består av två olika typer av underliggande autocalls; Phoenix Autocall och Autocall Uncapped. Det konstaterades att metoden med en ekvi-korrelationsmatris hade i en majoritet av fallen de minsta avvikelserna från motpartspriset. Ett annat resultat var att de implicita korrelationsmatriserna presterade bättre på de optionskorgar med Autocall Uncapped än Phoenix Autocall som underliggande. Ett intressant ämne för vidare forskning är att undersöka andra marknader men också att studera metoderna när mer än en marknad granskas.

Contents

Chapter 1

1   Introduction

This chapter begins with an introduction to the background and previous research of the problem. In the following sections the problem is formulated, followed by the purpose and research questions of this thesis. The last section gives an review of the following chapters.

1.1   Regulatory framework

The regulatory agencies implement financial regulations which require financial institutions to follow certain requirements, restrictions and guidelines. These regulations are implemented in order to ensure market transparency and comparability of financial statements and thus an efficient functioning of the European capital market and of the internal market.

The International Financial Reporting Standards (IFRS) are accounting standards issued by IFRS Foundation and the International Accounting Standards Board (IASB) to provide a global framework for business affairs so that company accounts are understandable and comparable across international boundaries. One of the standards, *IFRS 9 — Financial Instruments*, is a regulatory framework that addresses the accounting for financial instruments.

Included in IFRS 9 is the fair value hierarchy, which is used to categorize fair value measurements into three different levels. The IASB's definition of fair value is *"the price that would be received to sell an asset or paid to transfer a liability in an orderly transaction between market participants at the measurement date"*. The most reliable evidence of fair value is a quoted price in an active market (Level 1). The lowest category (Level 3) has the highest level of disclosures and the inputs are not based on observable data. Assets are preferred to be classified at as high level as possible [1].

Some financial instruments have correlation as input in their valuation. Since correlation can not be directly observed in the market, these types of instruments will be classified as Level 3. One way to overcome this obstacle is to calculate the correlation that is implied from observable market data i.e. option prices. This would allow the asset to be classified in a lower level.

1.2   Previous research

There is plenty of research about different methods used to estimate correlation between assets. Correlation can not be observed directly in the market and therefore needs to be estimated from the context of a particular method. One way to estimate the correlation is under the objective measure from the time series of assets return, so called historical data. The methods using historical data are backward-looking and rely on the assumption

that the future is fairly similar to the past. A complicated case is when the number of correlation coefficients exceeds the time series length. This problem has been studied in different areas. The random matrix theory (RMT) studies the case where the dimension of assets correlation is large and the time series length is short in an asymptotic setting [2]. Another area is the study on developing regularization methods for sample covariance and correlation matrices, such as shrinkage technique [3] and bending [4].

Another way to estimate correlation, instead of entrusting a backward-looking method, is by using the risk neutral measure. This can be done by a current snapshot of the prices from the option market. Option prices may also contain supplementary information that is not enclosed by historical data. In contrary to historical data, option prices also contain information that are forward-looking and reflects the market participants expectation of the future price [5]. Correlation that is implied from option prices are called implied correlation and the modeling of this is a challenging task. This since the implied correlations are not constant over time, maturities and strikes, and also since the number of correlation coefficients to estimate increase with the size of a portfolio.

It has been detected that correlation implied from option prices have a high predictive power and is useful as a proxy for future realized correlations. When forecasting over a specific horizon, the future level of diversification in the portfolio can also be measured [6].

1.3   Problem formulation

The correlation of financial assets is of great value in finance. It can be used to predict the relationship between assets and is an important factor when assessing the asset's risk. Correlations are not observable in the markets though there are different methods that can be used to estimate it. The estimation of correlation can be made from historical asset returns or from option prices from the market.

When modeling correlations of specific variables, most methods only use observed values from the past as a relevant information set. This is not optimal since, as previous mentioned, historical data does not include important information from the current market and only builds on the belief that the future will be similarly to the past. Correlations can vary a lot over time, which makes it limiting and unreliable to use historical correlations for future predictions among assets. The other mentioned approach is to estimate correlations from current option prices, so called implied correlations. The option prices reflect the market's expectation of the future prices and contain additional information which is left out in historical data.

Correlation estimates are used when pricing a basket of stocks and an inaccurate correlation will lead to an incorrect valuation that does not coincide with counterparty valuation. Therefore this thesis is performed on behalf of Handelsbanken Capital Markets, who requested a foundation for different correlation methods that can be used when pricing baskets. For this reason, three different methods will be compared and evaluated in order to find a suitable and more accurate valuation method. The performance of the three methods

2

can be measured through pricing basket options and then compare those to counterparty prices.

1.4 Purpose

The aim of this study is to model correlation between assets based on observable market data.

1.5 Research questions

1. What are the pros and cons using a method built on market prices to calculate the correlation of the components in an index?

2. How can the problem with an invalid correlation matrix be avoided?

3. Does one of the tested methods outperform the other?

1.6 Outline

Chapter 2, provides an introduction to financial terms used in this thesis.

Chapter 3, covers the theoretical background that is necessary for this thesis. The theory for option pricing is explained. Then information about the process of establishing a realized and implied correlation matrix are provided.

Chapter 4, introduces three methods for calculating implied correlation.

Chapter 5, provides a general method for the three correlation models and then a more detailed explanation of every method is described. Also, the process of pricing a basket option is lightly explained.

Chapter 6, covers all the data used in this thesis. Information about financial market data, delimitations and tools are presented.

Chapter 7, shows the results of the tested methods. A comparison of heat maps on the tested correlation models and tables of price deviation from counterparty prices are illustrated.

Chapter 8, presents an analysis and discussion of the result.

Chapter 9, answers the research questions and summarizes this thesis. An introduction to further research is given.

Chapter 2

2 Terminology

This chapter gives a short introduction to relevant financial terms used in this thesis.

Derivative: financial security whose value is derived from or depends on the value of other underlying variables.

Underlying asset: financial instrument, such as futures, stocks, a index etc., that the derivative's price is based on.

Call option: contract between two parties which gives the buyer the right, but not the obligation, to buy a specific underlying asset at a certain price in the future.

Put option: contract between two parties which gives the buyer the right, but not the obligation, to sell a specific underlying asset at a certain price in the future.

European option: can be exercised only at a specific predetermined maturity, T

American option: can be exercised at any trading day on or before maturity, T.

Strike price, K: the fixed price at which a specific derivative contract, mostly stock and index options, can be exercised.

Spot price, S_0: the current price of a security at which it can be bought or sold at a particular place and time.

Future contract: a legal agreement to buy or sell something at a predetermined price at a specified time in the future.

At-The-Money (ATM): when the strike price of an option is the same as the current price of the underlying asset.

In-The-Money (ITM): for a call option it is when the strike price is below the current trading price and for a put option it is when the strike price is above the spot price.

Out-of-The-Money (OTM): when the strike price of a call/put option is above/below the trading price of the underlying security.

End-of-Day (EOD): the final closing price of a stock at the end on the day, it is when the stock market concludes its trading activity for the day.

Stock index: measurement of a section of the stock market that is computed from prices of selected stocks, usually a weighted average. A index is used as a tool that describes the market.

Structured products: investment strategy consisting of a forward or a future and/or an option, often with an index or stocks as underlying instruments. Structured products are often very complex and comes with high level of risk.

Autocallable products: a subcategory within the structured products, also called auto-calls. An autocall is linked to an obligation and a derivative.

Option pricing: different methods are used depending on type of the option. European options are priced using Black-Scholes model while American options can be priced using a numerical method.

Basket option: a type of financial derivative. The underlying assets is a group of stocks, commodities, indices etc.

Dividends: a payment, usually as a distribution of earnings, made by a corporation to its shareholders. Dividends can be issued as cash payments, as shares of stocks or other property.

Risk-free interest rate: the theoretical rate of return of an investment with zero risk. It represents the interest an investor would expect from an risk-free investment over a given period of time.

Volatility, σ: in option pricing, it is the variable showing the extent to which the return of the underlying asset will fluctuate until maturity. Usually, the higher the volatility, the riskier the security.

Correlation: statistical measure that shows whether and how strongly two variables are related to each other.

Chapter 3

3 Theoretical background

This chapter provides a foundation of the relevant theory used in this thesis. The first sections introduces probability measures and their relation in correlation. The following sections describes the theory used for option pricing and how it can be used to calculate implied volatilities. In the last section the theory for calculating the realized correlation matrix is given.

3.1 Probability measure $\mathbb{P}$ & $\mathbb{Q}$

Consider a market under the filtrated probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where the four quantities are described as:

Ω - the sample space, i.e. is the set of all possible outcomes w.

$\mathcal{F}$ - a σ-algebra of Ω. To construct a σ-algebra, the introduction of an event A is needed. An event is a collection of outcomes and a subset of the sample space; $A \subseteq \Omega$. Further, $\mathcal{A}$ is a collection of subsets of Ω and $\mathcal{A}$ is called a σ-algebra if

    (i) $\emptyset \in \mathcal{A}$

    (ii) if $A \in \mathcal{A}$ then $A^c \in \mathcal{A}$, where A^c denotes the complement of A

    (iii) $\mathcal{A}$ is closed under finite countable unions and finite countable intersections [7].

$\mathbb{F}$ - a family $(\mathcal{F}_t)_{t \geq 0}$ of sub-σ-algebras of $\mathcal{F}$; meaning that for each t, $\mathcal{F}_t$ is a σ-algebra included in $\mathcal{F}$ and if $s \leq t$, $\mathcal{F}_s \subset \mathcal{F}_t$. Such a family $\mathbb{F}$ is called a filtration.

$\mathbb{P}$ - a probability measure defined on $\mathcal{F}$, denoted as the objective or actual measure.

The probability $\mathbb{P}$ is the market participant subjective probability under consideration. The risk neutral probability measure $\mathbb{Q}$ is another probability measure on $\mathcal{F}$ such that $\mathbb{Q}$ and the objective probability $\mathbb{P}$ are equal in the sense of measure; they assigns zero probability to the same events, i.e. they agree on the null space. This implies, by the Radon-Nikodym theorem, that there exists a non-negative random variable Z such that $\mathbb{E}_\mathbb{P}(Z) = 1$ and for all random variables X,

$$\mathbb{E}_\mathbb{Q}(X) = \mathbb{E}_\mathbb{P}(XZ), \quad \text{where} \quad Z = \frac{d\mathbb{Q}}{d\mathbb{P}}. \tag{1}$$

The random variable Z in Equation (1) is the so called Radon-Nikodym derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$ and it summarizes their relation. When going from the risk neutral measure to the objective measure, the transformation can be done by the change of measure technique.

If Z is the Radon-Nikodym derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$ the Radon-Nikodym process $(Z_t)_{0<t<T}$ is given by the Martingale

$$Z_t = \mathbb{E}[Z|\mathcal{F}_t], \quad 0 \leq t \leq T.$$

Passing in the inverse direction, from $\mathbb{P}$ to $\mathbb{Q}$, the Radon-Nikodym derivative $Z^{-1} = \frac{d\mathbb{P}}{d\mathbb{Q}}$ is required [8].

The existence of a risk neutral measure $\mathbb{Q}$ is essential in finance. Parameters extracted from option prices in the market, such as implied volatilities and implied correlations, are risk neutral values under the $\mathbb{Q}$ measure. The presence of a risk neutral measure is equivalent to the no-arbitrage condition. The absence of arbitrage is the fundamental and basic intuition when valuing assets. An arbitrage-free market implies that there should not be an opportunity to gain profit without exposure to risk.

3.2   Correlation

The payoff of an investment usually depends on the performance of multiple underlying instruments. Since the assets often have some degree of linear dependence it is important to be aware of the correlation between them. One must distinguish between the realized and the implied correlation of an asset.

The realized correlation, also called historical correlation, is a measure of dependences of assets during a period of time. The realized correlation, ρ, between two assets can be described by the relation

$$\rho_{X,Y}^{\mathbb{P}} = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var(X) Var(Y)}}} \tag{2}$$

where X and Y are historical log-returns of two assets. When Equation (2) is used to calculate the realized correlation of two series of data, it is important that the series have matching dates. This gets problematic when trying to measure correlation between assets in different markets due to e.g. public holidays. In order to obtain a more profound view of the dependence between assets than the linear relationship between them, copulas can be used.

The calculation of the implied correlation is cumbersome since the market for European options on pairs of underlying baskets is not liquid so one can not extract an implied correlation between the underlying from these prices. Instead, one approach is to use market quotes on an index option and each of the underlying assets composing the index to proxy an implied correlation between the components of the index.

The correlation coefficient, ρ, has an interval between $[-1, 1]$, where a perfect negative correlation is indicated by a correlation value equal to -1 and a perfect positive correlation is indicated by a correlation value equal to 1. A correlation of value zero signals that two variables move in a generally random manner comparatively. The pairwise correlation between two assets is based on e.g. the assets industries and market cap. The correlation

between multiple assets in a portfolio can be described by a correlation matrix, R.

Definition 3.1. A valid $n \times n$ correlation matrix must fulfill the following constraints [9]:

1. The correlation matrix is symmetric, $\rho(A,B)=\rho(B,A)$

2. The diagonal entities must be equal to one, $\rho(A,A)=1$

3. Non-diagonal entries are real numbers, $\rho \in [-1,1]$

4. The correlation matrix must be positive semi-definite (meaning that all eigenvalues are non-negative) otherwise it must be modified such that condition is met.

The variance of a portfolio, containing an index, consisting of n assets is given by

$$\sigma_{index}^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \rho_{i,j} w_i w_j \sigma_i \sigma_j \tag{3}$$

where
$\sigma_{index} =$ implied volatility of the index
$w_i, w_j =$ weight of asset i and j for $i,j = 1,\ldots,n$
$\sigma_i, \sigma_j =$ implied volatility of asset i and j for $i,j = 1,\ldots,n$
$\rho_{i,j} \quad =$ correlation coefficient between asset i and j for $i,j = 1,\ldots,n$
$\rho_{i,i} \quad = 1$ for $i = 1,\ldots,n$.

Suppose that a portfolio assumes that the relationship among assets is described by an equicorrelation matrix, i.e. all correlation coefficients that are non-diagonal, $\rho_{i,j}$ for $i \neq j$, are identical. Then all correlation coefficients $\rho_{i,j}$ for all $i \neq j$ can be set to a constant $\rho \in [-1,1]$. The relationship between assets in this portfolio is described by rewriting Equation (3) to

$$\sigma_{Index}^2 = \sum_{i=1}^{n} w_i^2 \sigma_i^2 + 2\rho \sum_{j=1}^{n-1}\sum_{i>j}^{n} w_i w_j \sigma_i \sigma_j. \tag{4}$$

The weight of an asset in the portfolio is determined as

$$w_i = \frac{P_i S_i}{\displaystyle\sum_{i=1}^{n} P_i S_i} \tag{5}$$

where P_i is the price of the i:th index component and S_i is the float-adjusted shares outstanding of the i:th index component. Typical for a stock index is that the weight of each stock is dependent on its market value, so called float-adjusted. This means that an index

only counts those shares that are accessible to investors, excluding shares held by government or other companies. By reconstructing Equation (4) and solving for ρ, the closed-form formula is given as

$$\rho_{index}^{\mathbb{Q}} = \frac{\sigma_{index}^2 - \sum_{i=1}^{n} w_i^2 \sigma_i^2}{2 \sum_{1 \leq i < j \leq n} w_i w_j \sigma_i \sigma_j}. \tag{6}$$

In Equation (6), ρ reflects the market-capitalized weighted average correlation of an index's assets [9]. The relation between realized and implied correlation can be expressed through the theory of *weighted average correlation matrices (WACM)* where the correlation coefficient, $\rho_{index}^{\mathbb{Q}}$, is set to

$$\rho_{index}^{\mathbb{Q}} = \rho_{index}^{\mathbb{P}} - \alpha(1 - \rho_{index}^{\mathbb{P}}), \tag{7}$$

where α relates to the realized and implied correlations of the index and is set to be a constant in the interval $(-1, 1)$.

If α is assumed to be negative, the implied correlations will be higher than the realized correlation and if α is positive, the opposite outcome holds. When the realized correlation is higher than the implied, an investor might want to sell correlation by e.g. selling a call option on the index and buy a portfolio of call options on the individual constituents of the index [10].

3.3 Autocallable products

New financial instruments are constantly being created by banks and investment firms to meet their customers' needs. A lot of these new instruments are categorized as structured products which is an investment strategy consisting of a forward or a future and/or an option, often with an index or stocks as underlying instruments. Structured products are often very complex and comes with high level of risk.

A subcategory within the structured products are the autocallable products, also called autocalls. An autocall is linked to an obligation and a derivative. The derivative is often a combination of one or several call options and a sold put option. The maturity is often five years but an autocall can also mature automatically prior to their scheduled maturity date. The autocalls automatically matures if certain predetermined market conditions are met with the underlying assets in consideration. Depending on the underlying asset's observed level, the outcome of an autocall will differ. If the observed level is at or above the upper barrier, seen as a trigger level, the autocall prematures. When the scheduled maturity is reached and a kick out has not occurred, the observed level can either be within the barrier levels or below. If the observed level is within the barriers the invested capital will be returned fully, but if the observed level exceeds the lower barrier the investor is unprotected to the underlying assets downside [11]. This is exemplified in Figure 1.
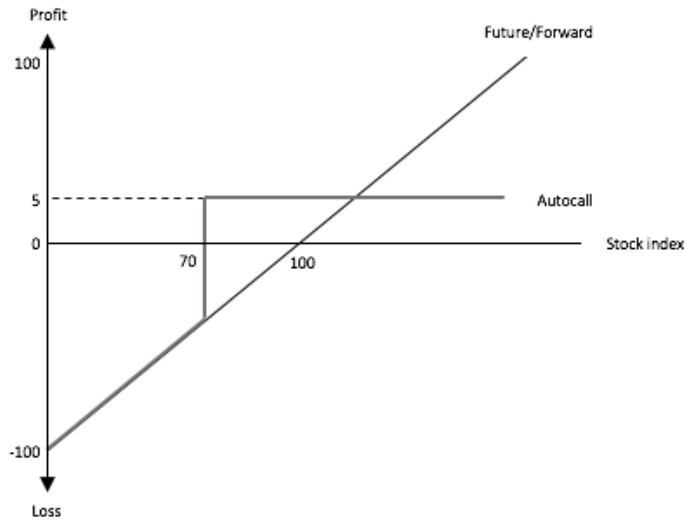
Figure 1: An example of an autocall with a risk return of 5 percent and a barrier level of 70 percent on the downside. This means that the stock index can loose 30 percent without the investor losing any money.

3.4 Black's model

Black's model, also refereed to as Black '76, is an adjusted version of the Black-Scholes option pricing model. A presentation of Black-Scholes model is given before continuing to Black '76. The simplest version of Black-Scholes model contains a risk-less asset and a risky asset. Certain assumptions are made when using the Black-Scholes model, such as:

- Used options are European and can therefore only be exercised at maturity.

- The risk-free rate and volatility of the underlying are constant and known.

- During the life of the option, no dividends are paid out.

- Movements in the market can not be predicted.

- When buying the option, there are no transaction costs.

- The return on the underlying are normally distributed.

Consider actively traded European call and put options at time 0 with prices C_0 and P_0 respectively. The spot price is S_0, strike price K and time to maturity is T. The discount

10

factor is denoted as $B_0 = e^{-rT}$ where r is the risk-free interest rate. Then the Black-Scholes model is given as

$$C_0 = S_0\Phi(d_1) - B_0K\phi(d_2)$$
$$P_0 = B_0K\Phi(-d_2) - S_0\Phi(-d_1)$$

where
$$d_1 = \frac{\log(S_0/(B_0K))}{\sigma\sqrt{T}} + \frac{\sigma\sqrt{T}}{2} \qquad \text{and} \qquad d_2 = d_1 - \sigma\sqrt{T}.$$

The cumulative distribution function of the standard normal distribution is denoted as Φ.

In contrast to Black-Scholes model the Black '76 model makes assumptions such that the expected change in futures price is zero and that the future prices are log-normally distributed. Another assumption is that the volatility is dependent of time.

Let the options be on a market index. Since the index does not pay any dividends, $S_0 = B_0G_0$ where G_0 is the future price of the index. Then B_0 can be derived from the put-call parity
$$C_0 - P_0 = S_0 - B_0K. \tag{8}$$

The zero rates can be calculated using $r = -\log(B_0)/T$. Multiple values for the zero rate can be derived using different prices on puts and calls for several strikes. From market the price of a call option, together with strike price and the future price can be inserted into Black's formula
$$C_0^B = B_0(G_0\Phi(d_1) - K\Phi(d_2)) \tag{9}$$

where
$$d_1 = \frac{\log(G_0/K)}{\sigma\sqrt{T}} + \frac{\sigma\sqrt{T}}{2} \qquad \text{and} \qquad d_2 = d_1 - \sigma\sqrt{T}.$$

The implied volatility, σ, can then be derived from back-solving for the value of the volatility in Equation (9) [12].

3.5 Finite difference method

The finite difference methods (FDM) are discretization methods. In general, FDMs are numerical methods where differential equations, $df(x)$, are solved by approximating them to difference equations, Δf. The approximation is denoted as

$$\frac{df(x)}{dx} \approx \frac{\Delta f}{\Delta x}.$$

The difference equations are afterward solved iteratively.

One FDM is the Crank-Nicolson Method [13] which is a combination of the implicit and the explicit method. The process of pricing an option can be divided into four steps:

1. *Discretize the Black-Scholes partial differential equation (PDE).*
   The Black-Scholes PDE is written as

   $$\frac{\delta f(t,S)}{\delta t} + rS\frac{\delta f(t,S)}{\delta S} + \frac{1}{2}\sigma^2 S^2 \frac{\delta^2 f(t,S)}{\delta S^2} = rf(t,S)$$

   where S is the asset's value, r is the risk-free interest rate and $f(t,S)$ is the unknown value of an option. The appropriate approximations to use to achieve the Crank-Nicolson FDM are the central approximation for $\frac{\delta f}{\delta t}$ and $\frac{\delta f}{\delta S}$, and also the standard approximation for $\frac{\delta^2 f}{\delta S^2}$. The approximations will consider node $f_{i-0.5,j}$, see Figure 2. The approximations are then substituted into the Black-Scholes PDE and the Crank-Nicolson FDM is given as

   $$-\bar{a}_j f_{i-1,j-1} + (1 - \bar{b}_j)f_{i-1,j} - \bar{c}_j f_{i-1,j+1} = \bar{a}_j f_{i,j-1} + (1 + \bar{b}_j)f_{i,j} + \bar{c}_j f_{i,j+1}$$

   where

   $$\bar{a}_j = \frac{\delta t}{4}(\sigma^2 j^2 - rj), \quad \bar{b}_j = -\frac{\delta t}{2}(\sigma^2 j^2 + r), \quad \bar{c}_j = \frac{\delta t}{4}(\sigma^2 j^2 + rj)$$

   and σ is the volatility.

2. *Specify a grid of underlying asset prices.*
   The aim here is to create a grid of future potential prices of the underlying asset. The time axis of the grid is divided into M equal periods from today to expiry, $j = 0, 1, \ldots, M$ and the underlying prices are divided into N equal levels, $i = 0, 1, \ldots, N$. The nodes in a grid are represented by indices i and j, where this method prices all three of the left side nodes based on the values of all three of the right side nodes.
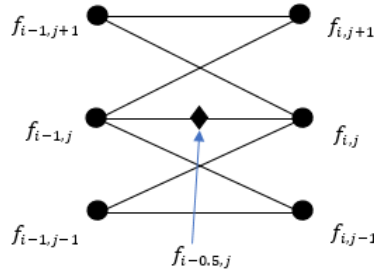

Figure 2: The grid for the Crank-Nicolson Method, each node represents price points.

3. *Calculate the payoff of the option at specific boundaries of the grid.*
   The boundary conditions are then calculated payoffs at each node on the boundary of

the grid. For the y-axis, the boundary condition is the payoff for S_{min} and the payoff for S_{max}. The boundary condition on the x-axis is the payoff at expiry.

4. *Determine the option price.*
   After the boundary conditions are specified the interior points can be computed. The backward iterative approach is used to work through time until the option price at grid nodes for today, $t = 0$, can be calculated.

3.6   Implied volatility of index options

Option prices reflect the market's expectations and can also be refereed to as the risk of a stock or a stock index. The amount of risk reflected in option prices can be expressed as the implied volatility. This makes the implied volatility one of the crucial factors when deciding the price of an option.

The implied volatility of an index option reflects the future volatility of that index's price return. An index option's volatility is affected by the index components individual volatilities and the correlations of the components price returns.

The relationship between the implied volatilities of options on an index and the implied volatilities of a weighted portfolio of options containing the same components will be a measure of the market's expectation on the index components future correlation. This measure is denoted as the implied correlation of the index and is a measure of comovement that display the estimated changes in the relative premium between index options and single-stock options [14].

Calculating the implied volatility of index options is a rather straight forward process since they are of European type. This means that there are no dividends or early exercise that needs to be taken into account. Hence, the Black-Scholes option pricing model can be used. There are various methods of the Black-Scholes model that can be applied depending on the type of the underlying instrument. In this thesis Black's model will be used.

3.7   Implied volatility of single stock options

A single stock option's implied volatility reflects the market's expectation on the future volatility of its price return. Calculating the implied volatilities of single stock options is more difficult than calculating the implied volatilities of index options since they are of American type. This means that early exercise and dividends needs to be taken into account. For those cases the standard Black-Scholes model no longer holds.

There are various numerical methods that can be applied for valuing options of American type. The methods all have their advantages and disadvantages. One approach is the Binomial Model which provides a generalizable method for the valuation of options. The method is both mathematically more simple and accurate than the Black-Scholes formula. Unfortunately the process is rather slow. The similar Trinomial Model approaches an accurate value faster than its binomial counterpart due to the use of a three-pronged path.

The Finite Difference Method (FDM) reminds of the Binomial Model but is more flexible since it can be applied to numerous exotic options. Among the FDMs, the Crank-Nicolson Method has the highest precision with a second order accuracy, in both the x- and t-direction. The method is also unconditionally stable and has an advantage when time-accurate solutions are important. For this reason the Crank-Nicolson Method is an appropriate valuation method for American options. Even though most methods for pricing American options are numerical there exist analytical models such as the Barone-Adesi and Whaley Model which is an accurate approximation method [15] [16].

3.8 Constructing a volatility surface

It is necessary to use different models when calculating the volatility since the stock options and index options are of different types. The implied volatility of single stock options which are of type American are calculated using a *Finite difference method (FDM)* and the index option of European type are derived from Black's model.

The derived values of implied volatilities are then plotted in a volatility surface using interpolation among the calculated values. The shape of the volatility may appear different depending on the options they represent. Near-term equity options implied volatility might get a more U-shaped appearance in a plot, as a typical smile. The information a volatility smile gives is that the demand for options that are ITM and OTM is greater than ATM options. Another pattern that is even more common is the reverse skew, see Figure 3. This shape goes by the name volatility smirk and typically appears for options with longer maturity, such as stock index options [17]. In general, the implied volatility smirk suggests that ITM calls and OTM puts (the dots above the ATM line in Figure 3) are more expensive in comparison to OTM calls and ITM puts (the dots below the ATM line in Figure 3) [18]. The volatility smirk occurs when the implied volatility is higher on options at lower strikes than the implied volatility on options at higher strikes.

When plotting the volatilities in a 3D-graph, as in Figure 5, it is easy to estimate the implied volatilities for different strikes and maturities.
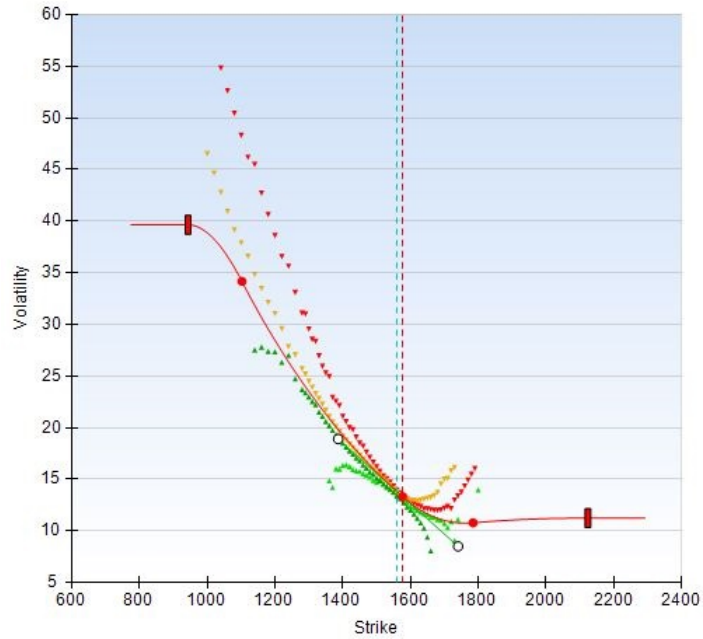
Figure 3: Volatility surface of OMXS30. The darker green dots represent bid prices of put options while the lighter green represent bid prices of call options. The red and orange dots represents ask prices for put and call options respectively. The red solid line is the interpolated implied volatility and the red dashed line symbols ATM options. Below the ATM the puts is ITM while the calls are OTM. Above ATM it is the other way around.
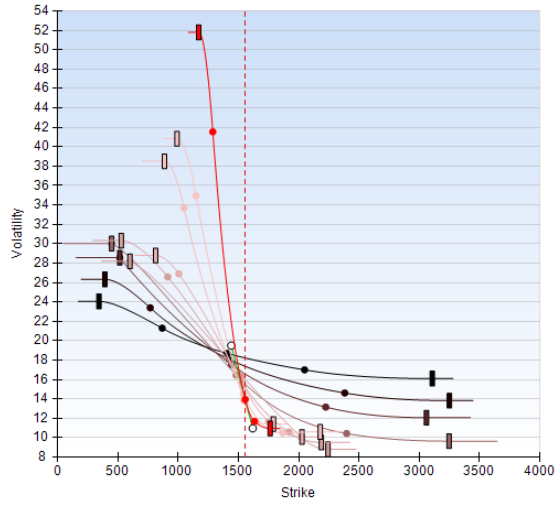
Figure 4: Volatility surface of OMXS30 for maturities up to 5 years. The volatility gets a more horizontal look as the maturities get longer.
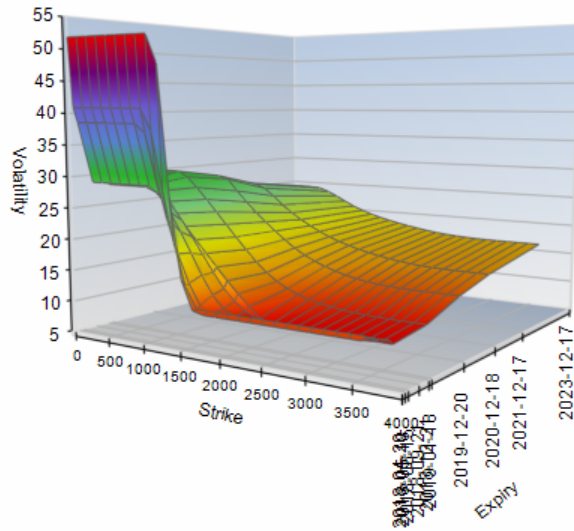
Figure 5: A 3D-graph of OMXS30.

3.9 Realized correlation calculations

It is a straight forward process to calculate the realized correlation, where the only needed input is historical stock prices. From the stock prices the historical log-return can easily be derived. Each sample of log-returns is used against all the other log-returns in Equation (2) to calculate the correlation between the assets in a portfolio. The correlation coefficients are the elements in the realized correlation matrix, $R^{\mathbb{P}}$.

Chapter 4

4 Implied correlation methods

In this chapter is the theory behind the three selected methods described.

4.1 Method 1 — Correlation with bump

One strategy for achieving a correlation matrix is to use a relation between the implied and realized correlation. The relation is given by Equation (7), where $\rho_{index}^{\mathbb{Q}}$ can be rewritten as $\rho_{i,j}^{\mathbb{Q}}$ for each pairwise correlation. The equation is given as

$$\rho_{i,j}^{\mathbb{Q}} = \rho_{i,j}^{\mathbb{P}} - \alpha(1 - \rho_{i,j}^{\mathbb{P}}) \tag{10}$$

where α in this method is predetermined and set to be a constant in the closed interval $(-1, 1)$. Each element $\rho_{i,j}^{\mathbb{P}}$ represents the pairwise correlation between asset i and asset j. In the case when $i = j$ then $\rho_{i,i} = 1$ since the correlation between an asset and itself is naturally one. The relationship can be written in matrix form as

$$R^{\mathbb{Q}} = R^{\mathbb{P}} - \alpha \times (U - R^{\mathbb{P}}) \tag{11}$$

where U is a $n \times n$ matrix with all elements equal to one and $R^{\mathbb{P}}$ is a correlation matrix under the objective measure. After $R^{\mathbb{P}}$ is generated and α is chosen, Equation (11) can be used to obtain $R^{\mathbb{Q}}$.

4.2 Method 2 — CBOE implied correlation index

Chicago Board Options Exchange, CBOE, has developed a method for calculating implied correlation index by using the S&P 500 Index (SPX), which is called CBOE S&P 500 Implied Correlation Index. CBOE daily disseminates values for the CBOE S&P 500 Implied Correlation Index. The daily values are market-based estimates of the expected average correlation of the components included in the SPX [14]. The implied correlation is calculated according to the theory of an equicorrelation matrix in Section 3.2 where

$$\rho_{index}^{\mathbb{Q}} = \frac{\sigma_{index}^2 - \sum_{i=1}^{n} w_i^2 \sigma_i^2}{2 \sum_{1 \leq i < j \leq n} w_i w_j \sigma_i \sigma_j}.$$

The calculated $\rho_{index}^{\mathbb{Q}}$ is then used to generate the implied correlation matrix $R_{CBOE}^{\mathbb{Q}}$.

4.3   Method 3 — Buss & Vilkov's Method

In the method described by Buss & Vilkov [9], the correlation matrix is not an equicorrelation matrix and the correlation coefficients between the assets will vary.

Let $\sigma_{Port}^{\mathbb{Q}}$ be the implied volatility of a portfolio, then the portfolio's variance is calculated by

$$(\sigma_{Port}^{\mathbb{Q}})^2 = W \cdot V^{\mathbb{Q}} \cdot R^{\mathbb{Q}} \cdot V^{\mathbb{Q}} \cdot W' \tag{12}$$

where

$$W = \begin{bmatrix} w_1 & \cdots & w_n \end{bmatrix} \tag{13}$$

$$V^{\mathbb{Q}} = \begin{bmatrix} \sigma_1^{\mathbb{Q}} & 0 & \cdots & 0 \\ 0 & \sigma_2^{\mathbb{Q}} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_n^{\mathbb{Q}} \end{bmatrix} \tag{14}$$

$$R^{\mathbb{Q}} = \begin{bmatrix} 1 & \rho_{2,1}^{\mathbb{Q}} & \cdots & \rho_{n-1,1}^{\mathbb{Q}} & \rho_{n,1}^{\mathbb{Q}} \\ \rho_{2,1}^{\mathbb{Q}} & 1 & \cdots & \rho_{n-1,2}^{\mathbb{Q}} & \rho_{n,2}^{\mathbb{Q}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{n-1,1}^{\mathbb{Q}} & \rho_{n-1,2}^{\mathbb{Q}} & \cdots & 1 & \rho_{n-1,n}^{\mathbb{Q}} \\ \rho_{n,1}^{\mathbb{Q}} & \rho_{n,2}^{\mathbb{Q}} & \cdots & \rho_{n-1,n}^{\mathbb{Q}} & 1 \end{bmatrix}.$$

The elements, w_i for $i = 1, \ldots, n$, in Equation (13) is calculated according to Equation (5). The matrix $V^{\mathbb{Q}}$ consists of the stocks' individual implied volatilities. The theory of WACM is also used here to extract a implied correlation matrix $R^{\mathbb{Q}}$ from $R^{\mathbb{P}}$, as in Equation (11), where $\alpha \in (-1, 0]$.

By inserting Equation (11) into Equation (12), the following equation for α is obtained

$$\alpha = -\frac{(\sigma_{Port}^{\mathbb{Q}})^2 - W \cdot V^{\mathbb{Q}} \cdot R^{\mathbb{P}} \cdot V^{\mathbb{Q}} \cdot W'}{W \cdot V^{\mathbb{Q}} \cdot (U - R^{\mathbb{P}}) \cdot V^{\mathbb{Q}} \cdot W'}. \tag{15}$$

When α is found, $R^{\mathbb{Q}}$ can be derived using Equation (11).

If $U - R^{\mathbb{P}} \geq 0$ then $\alpha > 0$ only when $\sigma_{Port}^{\mathbb{P}} > \sigma_{Port}^{\mathbb{Q}}$. This may cause an invalid correlation matrix ($R^{\mathbb{Q}}$) when $\sigma_{Port}^{\mathbb{P}} < \sigma_{Port}^{\mathbb{Q}}$. This problem can be resolved using the so called New algorithm described below [19].

Let A and B be two valid correlation matrices of dimension $n \times n$, then C is also a valid correlation matrix of the same dimension such that

$$C = w \times A + (1 - w) \times B \tag{16}$$

where w is a weight in the interval [0,1].

Let U and L be the valid upper and lower bound of equicorrelation matrices with entries 1 and $-1/(n-1)$ respective.
By rearranging Equation (16) and also replacing C and B to $R^{\mathbb{Q}}$ and $R^{\mathbb{P}}$ respectively, then equation

$$R^{\mathbb{Q}} = R^{\mathbb{P}} + w \times (A - R^{\mathbb{P}}). \tag{17}$$

is obtained. Inserting Equation (17) in Equation (12) gives

$$w = \frac{(\sigma_{Port}^{\mathbb{Q}})^2 - (\sigma_{Port}^{\mathbb{P}})^2}{W \cdot V^{\mathbb{Q}} \cdot (A - R^{\mathbb{P}}) \cdot V^{\mathbb{Q}} \cdot W'} \tag{18}$$

where

$$(\sigma_{Port}^{\mathbb{P}})^2 = W \cdot V^{\mathbb{Q}} \cdot R^{\mathbb{P}} \cdot V^{\mathbb{Q}} \cdot W' \tag{19}$$

is the implied volatility of the portfolio obtained from $R^{\mathbb{P}}$.
The following algorithm can be used to achieve an implied correlation matrix from implied correlation index.

1. Calculate the volatility of the portfolio ($\sigma_{Port}^{\mathbb{P}}$) using Equation (19).

2. Select the boundary matrix

   (a) If $\sigma_{Port}^{\mathbb{P}} > \sigma_{Port}^{\mathbb{Q}}$ then use the lower bound i.e $A = L$.
   (b) If $\sigma_{Port}^{\mathbb{P}} \leq \sigma_{Port}^{\mathbb{Q}}$ then use the upper bound i.e $A = U$.

3. Calculate w using Equation (18) with the selected boundary matrix from previous step

   (a) If $\sigma_{Port}^{\mathbb{P}} > \sigma_{Port}^{\mathbb{Q}}$ then

   $$w = \frac{(\sigma_{Port}^{\mathbb{Q}})^2 - (\sigma_{Port}^{\mathbb{P}})^2}{W \cdot V^{\mathbb{Q}} \cdot (L - R^{\mathbb{P}}) \cdot V^{\mathbb{Q}} \cdot W'}.$$

   (b) If $\sigma_{Port}^{\mathbb{P}} \leq \sigma_{Port}^{\mathbb{Q}}$ then

   $$w = \frac{(\sigma_{Port}^{\mathbb{Q}})^2 - (\sigma_{Port}^{\mathbb{P}})^2}{W \cdot V^{\mathbb{Q}} \cdot (U - R^{\mathbb{P}}) \cdot V^{\mathbb{Q}} \cdot W'}.$$

4. Calculate $R^{\mathbb{Q}}$ from Equation (17).

Chapter 5

5 Methodology

This chapter explains how the chosen methods are implemented and how their performance are tested in the process of pricing basket options.

5.1 Correlation methods

Three different methods are tested (1) Correlation with a bump, (2) CBOE's implied correlation method and (3) Buss & Vilkov's method. Method 1 is the simplest of the tested methods where the relation between the realized and implied correlation is fixed using a set of predetermined alphas. Method 2 is an equicorrelated matrix where all the components have the same correlation coefficient. The method can be a good choice when there are difficulties to obtain liquid option prices on the underlying. Method 3 is similar to method 1 but instead of a predetermined alpha, the coefficient is derived in more complex manner.

The implied correlation matrices are calculated for two different days to be able to compare the results. The correlation matrices are then used to price a number of selected basket options. The pricing process will be done for each of the three methods and repeated during red days. For method 1, only one of the α will be used in the process of pricing basket options since the other two correlation matrices with bump only affects the result by a small shift up or down. Due to the time consuming process of changing the correlation matrices every day and since the correlation can be approximated to be the same between a short period of time, the correlation matrices from only one of the two testing days will be used during the four day pricing process.

5.1.1 Method 1 — Correlation with bump

This methods sets α in Equation (11) to a constant. The correlation matrix $R^{\mathbb{P}}$ contains pairwise correlation for all the stocks in the basket. In the cases where $R^{\mathbb{P}}$ is not a valid matrix, according to Definition 3.1, the MATLAB function *nearestspd* is used. The function *nearestspd* finds the nearest matrix with the properties of being symmetric and positive definite.

The following steps are common for the methods in Section 4.2 and 4.3.

Step 1: The tracking basket.
When finding a tracking basket for an index with few constituents e.g. OMXS30 all the components can be used but when dealing with a larger index consisting of a large amount of components e.g. S&P500 it is preferred to only use a fraction of the components. The reason is because a index with a large amount of components will lead to a massive correlation matrix, that will be quite cumbersome to calculate. The selection of stocks can

be done in different ways; one way is to select the highest market capitalization ranked stocks in the index and another way is to use multiple linear regression. When using the second method the stocks are ranked in their explanatory order and the stocks with highest rank are selected. Both methods are computational effective and can not take non-linear constraints into consideration, i.e. transaction costs.

Step 2: Calculate the implied volatility of each stock option.
The implied volatility of the stock options chosen in the tracking basket are calculated using the theory described in Sections 3.7 and 3.8.

Step 3: Calculate the implied volatility of the index.
The implied volatility of the index option, $\sigma^{\mathbb{Q}}_{Index}$, is calculated from the volatility surface as described in Sections 3.6 and 3.8.

Step 4: Calculate the capitalization weights for the components in the stock index.
The weight of an index component, w_i, is calculated by Equation (5). For every i:th component the price and the float-adjusted shares are collected to be used in the calculations of w_i.

Recall that the implied correlation is calculated with the individual stock weights. Hence, it is important to know that when dealing with a large portfolio the used weights are determined relative to the capitalization of the tracking basket and not on the index basket itself.

5.1.2 Method 2 — CBOE implied correlation index

When *Step 1-4* is done, all the necessary input to be used in the CBOE method is achieved.

Step 5: Calculate the implied correlation matrix.
Use the theory in Section 4.2 to calculate the average implied correlation, $\rho^{\mathbb{Q}}$. When $\rho^{\mathbb{Q}}$ is obtained, the implied correlation matrix is generated as an equicorrelation matrix. The implied correlation matrix for the CBOE method has can be seen as

$$
R^{\mathbb{Q}}_{CBOE} = \begin{bmatrix} 1 & \rho^{\mathbb{Q}} & \cdots & \rho^{\mathbb{Q}} & \rho^{\mathbb{Q}} \\ \rho^{\mathbb{Q}} & 1 & \cdots & \rho^{\mathbb{Q}} & \rho^{\mathbb{Q}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho^{\mathbb{Q}} & \rho^{\mathbb{Q}} & \cdots & 1 & \rho^{\mathbb{Q}} \\ \rho^{\mathbb{Q}} & \rho^{\mathbb{Q}} & \cdots & \rho^{\mathbb{Q}} & 1 \end{bmatrix}.
\tag{20}
$$

In Equation (20) the correlation coefficients are equal and reflect the market-capitalization weighted average correlation of the index components.

5.1.3 Method 3 — Buss & Vilkov's Method

Calculations from *Step 1-4* is already attained, continuing to the next step.

Step 5: Calculate the implied correlation matrix.
The theory in Section 4.3 is used to obtain the implied correlation matrix, $R^{\mathbb{Q}}$, and starts with the below listed constructions and calculations:

- A diagonal matrix $V^{\mathbb{Q}}$, as in Equation (14), is generated by the assets individual implied volatilities which are calculated in *Step 2*.

- The calculated weights from *Step 4* forms a vector W, from Equation (13), with dimension $1 \times n$.

- Generating a matrix U, with the dimension $n \times n$, where all elements are equal to 1.

The above listing and the calculation of *Step 3* is used to obtain an α from Equation (15). For the calculated α to be valid in this method, it is critical that the α lies within its required boundaries. Fortunately there is a way to handle the invalidity. Depending on the outcome of α, the following algorithm is selected:

(i) If α is valid, that is $\alpha \in (-1, 0]$: the process ends here and α is inserted into Equation (11). The searched $R^{\mathbb{Q}}_{BV}$ is obtained and an inspection is done to see if it satisfy Definition 3.1.

(ii) If α is not valid, that is $\alpha \notin (-1, 0]$: the extended process below is needed to overcome the invalidity.

The constant w from Equation (18) demands computation of $\sigma^{\mathbb{P}}_{Index}$ in Equation (19). A comparison between $\sigma^{\mathbb{P}}_{Index}$ and $\sigma^{\mathbb{Q}}_{Index}$ will decide whether a lower or an upper boundary matrix will replace matrix A in Equation (18). The selected boundary matrix will adjust $R^{\mathbb{P}}$ up towards U or down towards L to obtain a valid implied correlation matrix $R^{\mathbb{Q}}_{NA}$. A verification of validity, before inserting w into Equation (17), is whether w lies in the interval $[0, 1]$. When that is done and the implied correlation matrix is obtained, the last examination is to check if the matrix fulfills the constraints of Definition 3.1.

5.2 Pricing basket options

There are numerous types of basket options and depending on the underlying instruments the valuation model differs. As mentioned in the Theoretical background chapter, some of the most complex financial instruments are structured products known as autocalls. In this thesis, basket options with two types of underlying autocalls are priced with the implied correlation matrices to test the correlation models performances. The autocalls are of types Autocall Uncapped and Phoenix Autocall. The methods are chosen since their valuation

model takes the correlation of the underlying assets into account and depending on the chosen correlation matrix the outcome might differ between the two option types. This thesis does not attach importance on the structure of the autocalls itself since they are only used as a tool to evaluate and compare the correlation matrices, the reader is encouraged to search for further readings within the subject.

The basket options are chosen such that at least two of the autocalls underlying instruments are included in OMXS30. Then the implied correlation matrices are used to derive a theoretical price of the basket option which can be compared to the counterparty mid price of the same basket option. If a basket option expires between two testing days, it has been removed from the data without being replaced. Also, if a price deviates too much from the counterparty price for all the methods, the basket is removed. The price calculations with the three test correlation matrices are done using a method that is already implemented in the bank.

Chapter 6

This chapter covers all the used data in this thesis. The first section gives detailed information about the extracted market data. The following sections gives an overview of data sources for financial data and what delimitations and tools are used in the study.

6 Data

In this thesis the following data is used:

- Historical EOD prices of OMXS30 stocks.

- Number of issued stocks.

For the three and 12 month put and call options the data below is collected:

- Prices on future contracts on OMXS30

- Spot prices on OMXS30 stocks

- Bid and ask prices

- Strike prices

Both the stock options and the OMXS30 index option are standardized, which means that their time to maturity, strike price and contract size are predetermined.

6.1 Data sources

There are numerous platforms from where financial data, such as real-time market data, news, fundamental data and analytics on different financial instruments can be extracted. Financial companies and institutions pays enormous amounts for this information. Two of the largest financial market data distributors are Thomson Reuters and Bloomberg. In Sweden e.g. Avanza or Nordnet can be used to access free financial data even if it often comes with an additional cost for real-time data access.

In this thesis, data is collected from Thomson Reuters and Nasdaq OMX Stockholm exchange.

6.2 Delimitations

This thesis focuses on correlation of stocks included in the Swedish stock index OMXS30 which consists of the 30 most actively traded stocks on the Stockholm Stock Exchange. The chosen stocks are traded on the same stock exchange, during the same opening hours, thus there is no need to take asynchronous data into consideration.

The time to maturity on the used index option and stock options is three and 12 months. This is because options that has up to a year to maturity provide a good balance between the liquidity of the underlying stocks and the stability of the implied volatilities.

6.3   Tools

An internal platform at Handelsbanken is used for collecting data from Nasdaq OMX and Reuters. Then MATLAB is used for the mathematical modeling.

Chapter 7

7 Results

This chapter covers the results of this thesis. The first sections shows the results of the three tested correlation models. In the last section the correlation models are compared in tables of deviation on prices of basket options.

7.1 Methods

The results of the calculated pairwise implied correlations of the underlying stocks in OMXS30 can be seen as heat maps in Figures 6-13. A heat map is a visual representation of data in the form of a map in which the data are represented as colors. This is an easy and effective way to illustrate and compare the differences between the tested methods. Two perfect positively correlated stocks are illustrated by a dark red square as seen in the diagonal of the heat maps below. If two stocks are perfect negatively correlated this is instead pictured with a dark blue square. A green square symbolizes zero correlation.

7.1.1 Method 1 — Correlation with bump

The method were tested for three fixed values of α. The chosen values are an estimated guess and set to -0.05, -0.10 and -0.15. The results of the pairwise correlation can be seen in the heat maps in Figures 6-11.
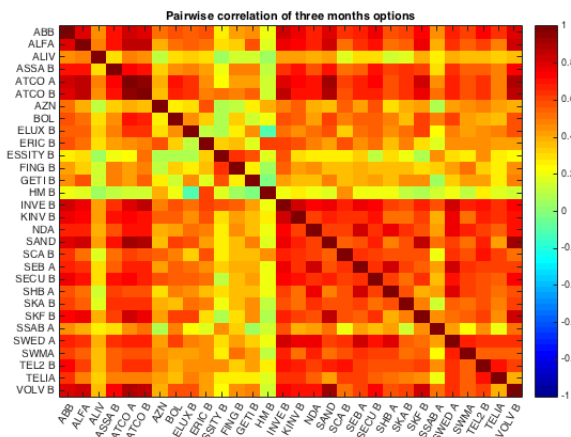

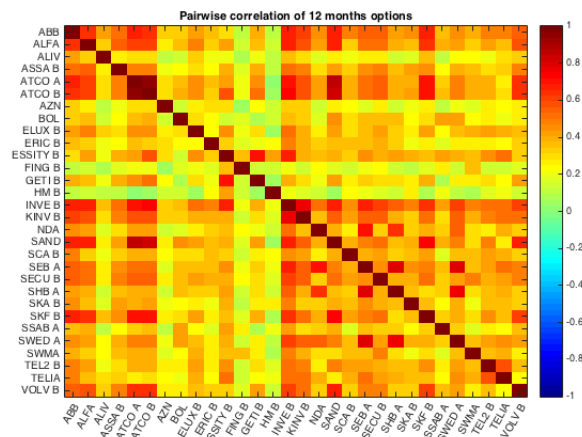Figure 6: Heat map using three month options where α is set to -0.05.

Figure 7: Heat map using 12 month options where α is set to -0.05.
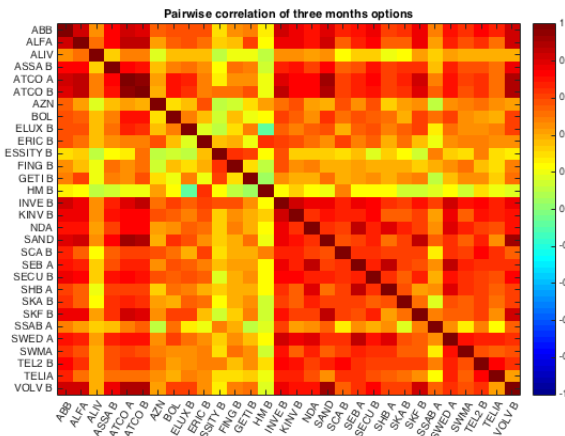
Figure 8: Heat map using three month options where α is set to -0.10.
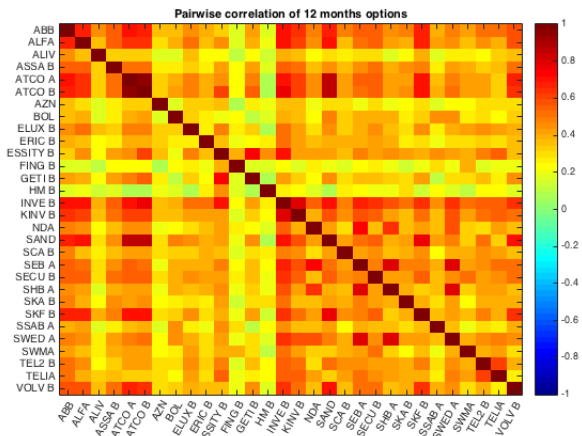
Figure 9: Heat map using 12 month options where α is set to -0.10.
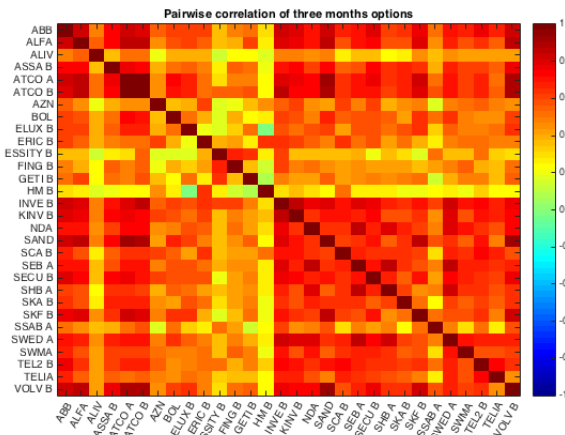
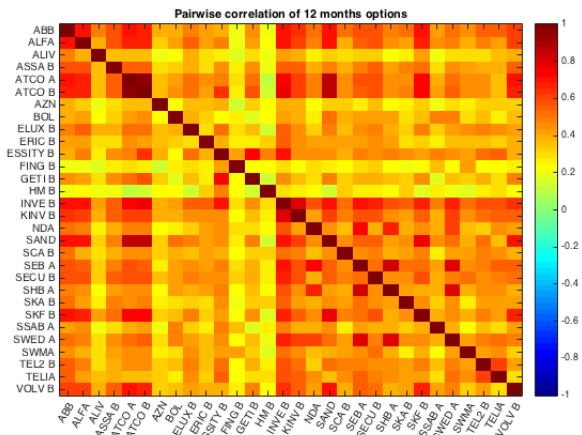Figure 10: Heat map using three month options where α is set to -0.15.

Figure 11: Heat map using 12 month options where α is set to -0.15.

7.1.2   Method 2 — CBOE implied correlation index

The average implied correlation were calculated using Equation (6) and for the three month options the returned value is 0.394. Inserting the value into Equation (20) gave the equicorrelated implied correlation matrix

$$R_{CBOE,3M}^{\mathbb{Q}} = \begin{bmatrix} 1 & 0.394 & \cdots & 0.394 & 0.394 \\ 0.394 & 1 & \cdots & 0.394 & 0.394 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0.394 & 0.394 & \cdots & 1 & 0.394 \\ 0.394 & 0.394 & \cdots & 0.394 & 1 \end{bmatrix}.$$

In the same way are the average implied correlation for 12 month options calculated, the returned value is 0.519 and displayed in the implied correlation matrix below

$$R_{CBOE,12M}^{\mathbb{Q}} = \begin{bmatrix} 1 & 0.519 & \cdots & 0.519 & 0.519 \\ 0.519 & 1 & \cdots & 0.519 & 0.519 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0.519 & 0.519 & \cdots & 1 & 0.519 \\ 0.519 & 0.519 & \cdots & 0.519 & 1 \end{bmatrix}.$$

The test was repeated during the second test day giving the values $\rho_{CBOE,3M} = 0.404$ and $\rho_{CBOE,12M} = 0.523$.

7.1.3 Method 3 — Buss & Vilkov's method

The calculation of α for the three month options resulted in an invalid α due to its positive value. From Section 5.1.3 (ii), it is known that an invalid α continues to the extended process of the New algorithm. The comparison between the values of $\sigma_{Index}^{\mathbb{P}} = 0.151$ and $\sigma_{Index}^{\mathbb{Q}} = 0.140$ resulted in a decision of letting a lower matrix L replace matrix A in Equation (18). The constant w were then calculated and the obtained value $w_{3M,1} = 0.139$ was valid. The implied correlation matrix were then extracted from Equation (11) and at last a check was done to see if the correlation matrix fulfilled Definition 3.1. The implied correlation matrix $R_{NA,3M}^{\mathbb{Q}}$ were valid and a heat map were constructed, seen in Figure 12.

The α calculated for the 12 month options came out valid with a value $\alpha_{12M,1} = -0.271$. Since α was valid, there was no need to apply the extended process of the New algorithm, instead Section 5.1.3 (i) was chosen. The implied correlation matrix, $R_{BV,12M}^{\mathbb{Q}}$, was obtained by inserting α into Equation (11). A validity check was done for the correlation matrix and all constraints of Definition 3.1 were satisfied. The heat map for $R_{BV,12M}^{\mathbb{Q}}$ was generated and can been seen in Figure 13. A similar result was received during the second test day where $w_{3M,2} = 0.124$ and $\alpha_{12M,2} = -0.267$ were calculated. Analyzing the validity of the generated implied correlations matrices for Day 2, the constraints of Definition 3.1 was fulfilled.
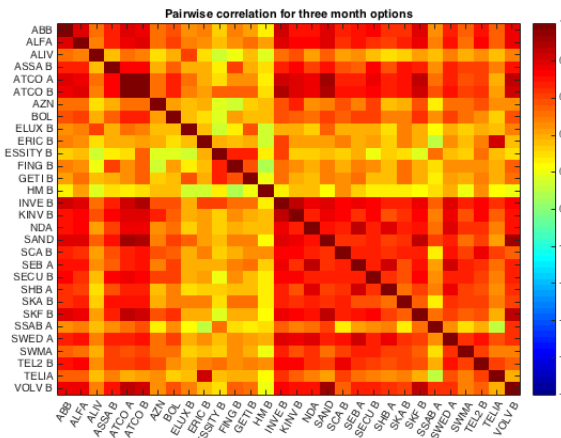
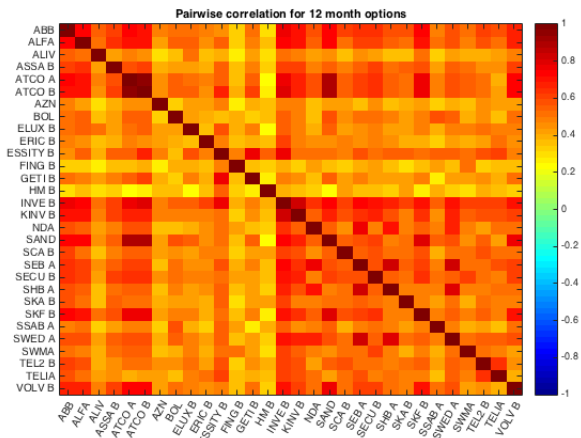Figure 12: Heat map from the first test day using three month options.

Figure 13: Heat map from the first test day using 12 month options.

7.1.4 Pricing basket options

The derived correlation matrices are used in the process of pricing basket options with underlying autocalls. The correlation matrices from Test 1 are used during each of the four days. From method 1, the pricing will only be made using the implied correlation matrix corresponding to where α was set to -0.10. The basket options are selected according to the theory in the Methodology chapter. During the first testing day, one basket option expired and was removed. A large price deviation from the counterparty price also led to a basket being excluded during the third testing day.

The result of the pricing process is represented as a deviation from the counterparty mid price. The test for Day 1 can be seen in Tables 1-4. In Table 1, basket 1-7 have Phoenix Autocall as underlying and in Table 2, baskets 8-12 have Autocall Uncapped as underlying. The following testing days are presented in Appendix A.

The results of the first day, Day 1, is represented in Table 1 and 2 in detail, where each individual basket's deviations and a resulting mean of each method are displayed. The following testing days, Day 2-4, are calculated in the same way and summarized in Table 3 and 4, where the mean of price deviations from each day and method are presented.

The scatter plot on Figure 14 visualize the mean of price deviations from Table 3 and Table 4. The x-axis represents the basket options with Phoenix Autocall as underlying and the y-axis represents the basket options with Autocall Uncapped as underlying. Each method is assigned a color and the marked dots in the figure represents a methods mean during a specific day.

Table 1: First day price deviation from counterparty price for Phoenix Autocalls.

Phoenix Autocall	Method 1	Method 2	Method 3
Basket 1	0.05	0.07	0.06
Basket 2	0.08	0.15	0.04
Basket 3	0.60	0.57	0.58
Basket 4	0.50	0.10	0.30
Basket 5	0.29	0.01	0.14
Basket 6	0.50	0.10	0.30
Basket 7	0.33	0.16	0.58
Mean	0.34	0.17	0.29

Table 2: First day price deviation from counterparty price for Autocall Uncapped.

Autocall Uncapped	Method 1	Method 2	Method 3
Basket 8	0.18	0.09	0.19
Basket 9	0.03	0.01	0.03
Basket 10	0.03	0.01	0.03
Basket 11	0.00	0.01	0.00
Basket 12	0.08	0.08	0.01
Mean	0.06	0.04	0.05

Table 3: Total mean of price deviation from counterparty price for Phoenix Autocalls.

Phoenix Autocall	Method 1	Method 2	Method 3
Day 1	0.34	0.17	0.29
Day 2	0.18	0.25	0.18
Day 3	0.28	0.18	0.17
Day 4	0.29	0.17	0.21
Total mean	0.27	0.19	0.21

Table 4: Total mean of price deviation from counterparty price for Autocall Uncapped.

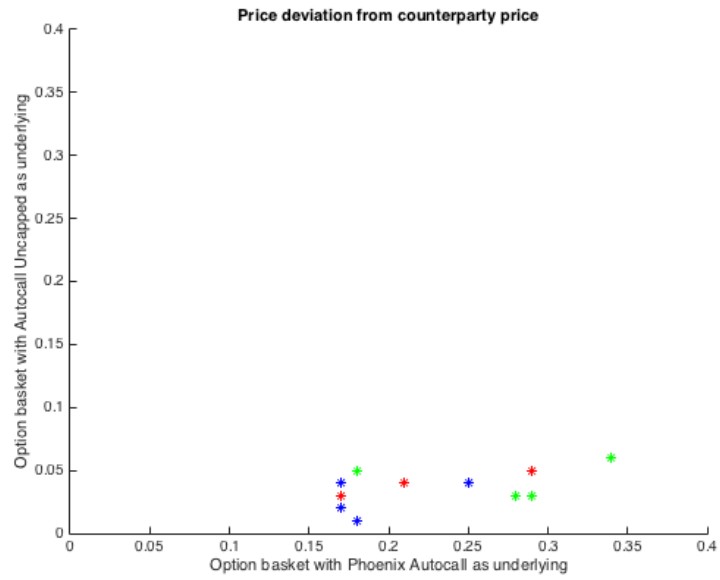| Autocall Uncapped	Method 1	Method 2	Method 3
Day 1	0.06	0.04	0.05
Day 2	0.05	0.04	0.05
Day 3	0.03	0.01	0.03
Day 4	0.03	0.02	0.04
Total mean	0.04	0.03	0.04

Figure 14: Scatter plot of the price deviation from the counterparty mid price during the four test days. Method 1 is illustrated by green dots, method 2 by blue dots and method 3 by red dots. Only three red dots are visible since during Day 2 method 1 and 3 have the same deviations.

Chapter 8

8   Discussion

This chapter contains an analysis of the results and possible sources of error.

8.1   Invalid correlation matrices

The realized correlation matrix for the OMXS30 stocks should, according to the Theoretical background chapter, be a positive semi-definite matrix but for all the testing methods a negative eigenvalue was derived. This might mean that the used data is not good enough. Since the implied volatilities are dependent on the maintenance of the volatility surfaces, this is a likely source of error. Since some of the most traded stocks in Sweden are used in this thesis, an assumption of surfaces being updated in the recent past is made. This does not exclude that the stocks volatility surfaces are updated on different days. Such problem can be avoided using market data on implied volatilities provided from e.g. Bloomberg.

To handle the breach of a negative eigenvalue the MATLAB function *nearestspd* is used.

8.2   Implied correlation matrices

For all the methods, Test 1 and Test 2, gave very similar results when extracting implied correlation matrices. In method 3 the boundaries of α are set, such that implied correlation always is higher than the realized correlation. This is a constraint up for debate since the case when the realized correlation is higher than the implied correlation should be a possible outcome in reality. Otherwise it would mean that the expected future correlation can not be lower than the recent history. Considering the three month case for method 3 a positive α was achieved, which lead to that the new algorithm was used to handle the invalidity. This approach makes the implied correlation higher than the realized correlation again. If instead, a positive α was considered as still valid, this would indicate that markets' expectation of the correlation will decrease in three months from the testing date. Since the realized correlation is based on historic data and this seem like a reasonable result. If considering the case of the three month correlation, the market has been shaky during the last three months. Within just a couple of days in February the Swedish volatility index (SIX) rapidly ascended and had not been that high since Brexit in June 2016, see Appendix B. During times of financial crises and when the market volatility is high, the correlation between stocks tends gets higher since people want to sell their assets independent of the stock company.

The heat maps for method 1 and method 3 shows that the implied correlation is higher in three months than in 12 months. This is not the case for method 2 where the implied correlation is higher in 12 months. When analyzing the 12 month implied correlation matrix for method 2, it was seen that the numerator was greater than the denominator. This since

the implied volatilities in the numerators second term of Equation (6) was lower for 12 month than three month.

In the heat maps for method 1 and 3, there is a dark red square which symbolized the almost perfect positive correlation between Atlas Copco's A (ATCO A) and B (ATCO B) stocks. For obvious reasons the correlation of two stocks of the same company should approximately be equal to one. Another stock that has a deviating pattern in the heat maps is Hennes & Mauritz (HM B). This could depend on the company's last months of media exposure which also affected the stock price to be very volatile. These patterns are not seen in method 2 since the correlation coefficient is an average of the implied correlation.

8.3 Pricing basket options

Due to difficulties in finding basket options where the correlation matrix is used in the valuation method, the amount of selected baskets is fewer than desired. In the results, the mean price deviation for the methods are given during four days. It would be preferred to have long history of price deviation from the counterparty price to be able to make a more solid examination of the correlation models. Another difficulty is the unawareness of how the counterparty sets their prices. In some cases the counterparty offers bid and ask price with a large spread since they might not want to make a trade and in other cases only a bid or ask price is offered.

When pricing the basket options, it is challenging to know how the correlation matrices will impact on the valuation model. From Table 1 it is given that method 2 has the lowest total mean during Day 1. Method 1 and method 3 have deviations with similar behavior where basket 3, 4, 6 and 7 have large deviations. The baskets 1 and 2 have the smallest deviations. Method 2 differs as much as 0.56 percentage point between the smallest and largest deviation. Worth noting is that basket 3 have the largest deviation for all the methods and if it was excluded, the mean deviation for method 2 would decrease significantly. In fact, it would nearly half its mean value from Day 1.

In Table 2 all the methods have a similar mean value but method 1 and 3 have large spreads between the smallest and largest deviation. In comparison to Table 1, the spreads are still small.

Looking at Table 3 and 4 an overview of the mean deviation during the testing days are given. For baskets containing Phoenix Autocall the total mean deviation is lowest for method 2 with a mean of 19%, closely followed by method 3 with a mean deviation of 21%. Method 3 is off with as much as 27%. The spread between the days are smaller for method 2 than for method 3, even though the two methods have close values in total mean.

For baskets containing Autocall Uncapped the total mean deviation is much lower. Method 2 is still closest to the counterparty price with a total deviation of 3%. The other methods have a slightly higher deviation of 4%. The daily spreads for baskets with Autocall Uncapped as underlying are very small in comparison to those with Phoenix Autocall as underlying.

Figure 14 gives an illustration of the resulting mean deviations for the two types of basket options. The dots are located close to the x-axis which implies that all the methods maintain a low deviation when pricing basket options with Autocall Uncapped as underlying.

It is also interesting to look at the mean spread between the two types of basket options. For baskets with Autocall Uncapped as underlying the mean deviation for the three methods are only 4% while the mean deviation for Phoenix Autocall as underlying are as much as 22%.

What can be seen from Tables 1 - 4 is that method 2 appear to be a better fit than the other methods for these specific baskets with underlying autocalls. This can be seen clearly in Figure 14. Not far-off comes method 3 with slightly higher deviations from counterparty prices. From the Theoretical chapter it is known that the implied correlation matrix used in method 2 is an equicorrelation matrix. This assumption is not realistic since pairwise correlation between different assets should not result in equivalent values in the real world. Which should make method 3 a more accurate method since it generates more realistic implied correlation matrices. However, it is known that the use of an equicorrelation matrix can overcome the obstacle of needing reliable option prices of bivariate options on all components in a portfolio. Yet, in this thesis the used stock options are liquid on the market so the lack of reliable option prices should not be a problem.

Chapter 9

This chapter presents the conclusions of this thesis and an introduction to further research.

9   Conclusions

The aim of this thesis was to model correlation between assets based on observable market data. Three different methods were tested to generate implied correlation matrices. Implied volatilities from the option market are used to create the implied correlation matrices. The benefit of extracting data from the option market is that it may contain additional information which is left out when using historical data, such as market's expectations of future prices. Also, correlation implied from option prices have a high predictive power and can be used as a proxy for future realized correlations. Unfortunately, there is not much research made with in the field.

When creating a correlation matrix it is crucial to ensure that it is valid. If the realized correlation matrix comes out invalid the MATLAB function *nearestspd* is used. This tool is used to approximate the correlation matrix to the nearest positive semi-definite matrix. In method 3, an additional condition was set to overcome the case of having an invalid implied correlation matrix. This condition is refereed as the New algorithm.

Basket options with two different underlying autocalls were priced with the different methods generated implied correlation matrices. The derived basket option prices were then compared to counterparty prices. Method 2 presented the lowest deviations from counterparty prices and have the smallest spread between the days. All methods seemed to fit basket options with Autocall Uncapped as underlying better than basket options with Phoenix Autocall as underlying. Also here method 2 outperformed the methods by suiting both of these basket options better, especially basket options with Phoenix Autocall as underlying.

9.1   Further Research

A further research on this thesis tested methods would be to consider the case when avoiding the harsh constraint on α, so that it is valid in the interval $[-1, 1]$ instead. Allowing α in the new interval will be more realistic since it is a possible outcome for the realized correlation to be greater than the implied correlation. However, it is important to examine if the implied correlation matrix is still valid.

Method 2 have been concluded as the better performing method with the chosen basket options with underlying autocalls. Therefore, it would be interesting to continue analyzing the model and how it operates on baskets options with other underlying autocalls.

Some delimitations have been made in this thesis that can be taken into consideration for further investigation. The simplification of only focusing on the Swedish market can be expanded to consider other markets independently, where e.g. S&P500 index can be

used on the American market or FTSE 100 Index on the UK market. Additionally, one can study the methods when more than one market is considered. Including assets from different markets, so called asynchronous data, can be problematic since it may involve countries with other trading hours and different public holidays. This may affect the pairwise correlations between assets to be wrongly estimated since the assets time series are inconsistent and will be mismatched.

For this thesis tested methods, a further research would be to examine their forecasting performance. As mentioned in the previous research, few studies have been done on the topic of using implied correlations from option prices as forecasts of future correlations.

For the interested ones, other tests can be done with implied correlation derived from another source, such as foreign exchange (FX) options. The Correlation implied from FX options have been studied and is found to be a useful tool for forecasting future currency correlations [20].

References

[1] International Accounting Standards Board. IFRS international financial reporting standards, 2018.

[2] Bai Z.D. Methodologies in spectral analysis of large dimensional random matrices. *A Review, Statistica Sinica*, 9:611, 1999.

[3] Ledoit O. and Wolf M. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of empirical finance*, 10(5):603–621, 2003.

[4] Bickel P. J. and Levina E. Regularized estimation of large covariance matrices. *The Annals of Statistics*, pages 199–227, 2008.

[5] Lopez J. and Walter C. Is implied correlation worth calculating? evidence from foreign exchange options and historical data. *Journal of Derivatives*, 7(3):65–81, 2000.

[6] Skintzi V. D. and Refenes A. N. Implied correlation index: A new measure of diversification. *Journal of Futures Markets*, 25(2):171–197, 2005.

[7] Brunnermeier M. K. Asset pricing under asymmetric information: Bubbles. *Crashes, Tech*, 2001.

[8] Bingham N. H. and Kiesel R. *Risk-neutral valuation: Pricing and hedging of financial derivatives.* Springer Science & Business Media, 2013.

[9] Buss A. and Vilkov G. Measuring equity risk with option-implied correlations. *The Review of Financial Studies*, 25(10):3113–3140, 2012.

[10] Bouzoubaa M. and Osseiran A. *Exotic options and hybrids: A guide to structuring, pricing and trading*, volume 505. John Wiley & Sons, 2010.

[11] Anjou M. and Oxenstierna G. *Specialister på värdepappersmarknaden.* Studentlitteratur, 2016.

[12] Hult H., Lindskog F., Hammarlid O., and Rehn C. J. *Risk and portfolio analysis: Principles and methods.* Springer Science & Business Media, 2012.

[13] Goddard consulting. Option pricing using the crank-nicolson finite difference method, 2018.

[14] Chicago Board Options Exchange. Cboe s&p 500 implied correlation index. Technical report, Working Paper, 2009.

[15] Recktenwald G. W. Finite-difference approximations to the heat equation. *Mechanical Engineering*, 10:1–27, 2004.

[16] Vidic I. *Numerical methods for option pricing.* PhD thesis, Informatica, 2012.

[17] Xing Y., Zhang X., and Zhao R. What does the individual option volatility smirk tell us about future equity returns? *Journal of Financial and Quantitative Analysis*, 45(3):641–662, 2010.

[18] Pathak R. and Mitra A. Predictability and predictors of volatility smirk: a study on index options. *Business: Theory and Practice*, 18:64, 2017.

[19] Numpacharoen K. and Numpacharoen N. Estimating realistic implied correlation matrix from option prices. *Journal of Mathematical Finance*, 3(04):401, 2013.

[20] Silyakova E. et al. Implied basket correlation dynamics. Technical report, Humboldt University, Collaborative Research Center 649, 2012.

Appendix A Price tables

Table 5: Day 2: Total mean of price deviation from counterparty price for Phoenix Autocalls.

	Method 1	Method 2	Method 3
Basket 1	0.13	0.14	0.13
Basket 2	0.06	0.33	0.20
Basket 3	0.57	0.54	0.54
Basket 4	0.27	0.17	0.04
Basket 5	0.02	0.27	0.12
Basket 6	0.04	0.02	0.04
Basket 7	-	-	-
Mean	0.18	0.25	0.18

Table 6: Day 2: Total mean of price deviation from counterparty price for Autocall Uncapped.

	Method 1	Method 2	Method 3
Basket 8	0.00	0.01	0.00
Basket 9	0.08	0.08	0.08
Basket 10	0.02	0.02	0.02
Basket 11	0.02	0.01	0.02
Basket 12	0.13	0.06	0.15
Mean	0.05	0.04	0.05

Table 7: Day 3: Total mean of price deviation from counterparty price for Phoenix Autocalls.

	Method 1	Method 2	Method 3
Basket 1	0.05	0.07	0.06
Basket 2	0.05	0.18	0.07
Basket 3	0.61	0.58	0.37
Basket 4	0.46	0.08	0.21
Basket 5	0.47	0.14	0.24
Basket 6	0.05	0.05	0.05
Basket 7	-	-	-
Mean	0.28	0.18	0.17

Table 8: Day 3: Total mean of price deviation from counterparty price for Autocall Uncapped.

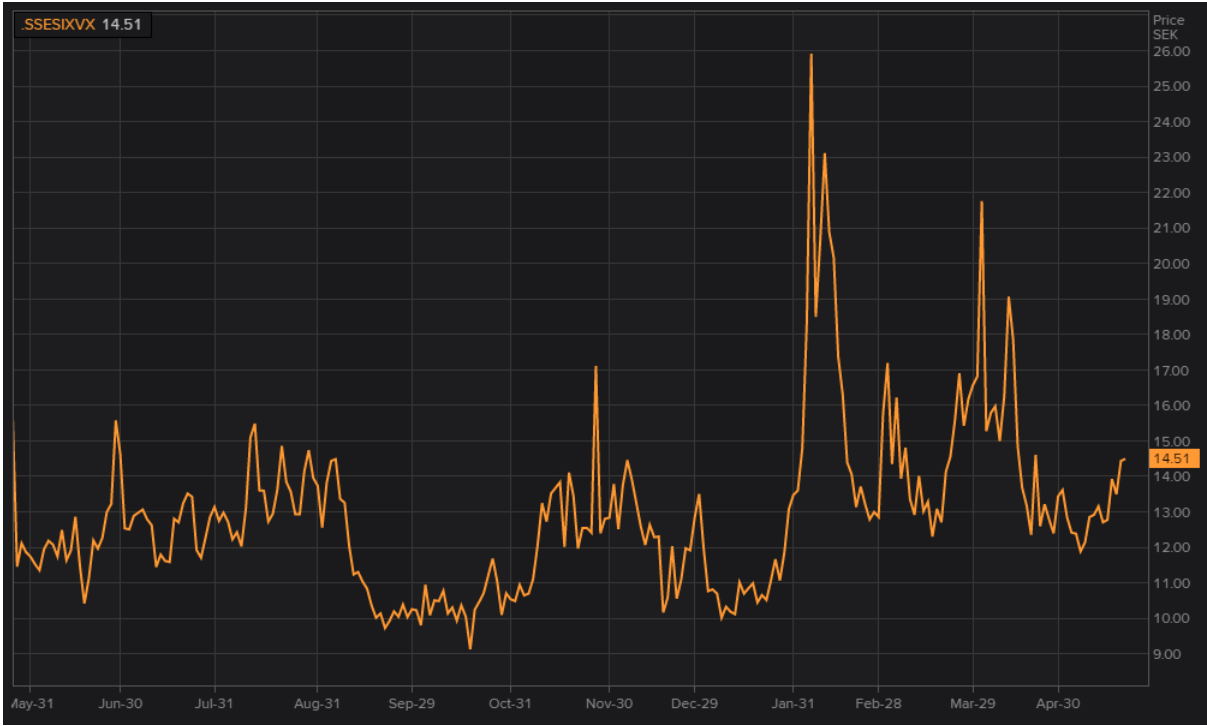|	Method 1	Method 2	Method 3
Basket 8	-	-	-
Basket 9	0.01	0.00	0.01
Basket 10	0.07	0.01	0.08
Basket 11	0.04	0.02	0.04
Basket 12	0.01	0.01	0.00
Mean	0.03	0.01	0.03

Table 9: Day 4: Total mean of price deviation from counterparty price for Phoenix Autocalls.

	Method 1	Method 2	Method 3
Basket 1	0.01	0.01	0.00
Basket 2	0.08	0.15	0.04
Basket 3	0.60	0.56	0.57
Basket 4	0.48	0.10	0.29
Basket 5	0.51	0.17	0.04
Basket 6	0.04	0.02	0.34
Basket 7	-	-	-
Mean	0.29	0.17	0.21

Table 10: Day 4: Total mean of price deviation from counterparty price for Autocall Uncapped.

	Method 1	Method 2	Method 3
Basket 8	-	-	-
Basket 9	0.06	0.05	0.06
Basket 10	0.00	0.01	0.01
Basket 11	0.06	0.00	0.07
Basket 12	0.01	0.01	0.01
Mean	0.03	0.02	0.04

Appendix B   SIX graph