MASTER'S DEGREE PROJECT IN FINANCIAL MATHEMATICS,
THIRD AND FOURTH CYCLE, 30 CREDITS
STOCKHOLM, SWEDEN 2023

Dispersion Trading: A Way to Hedge Vega Risk in Index Options

KTH Master Thesis Report

Albin Irell Fridlund
Johanna Heberlein

KTH ROYAL INSTITUTE OF TECHNOLOGY
Applied and Computational Mathematics

Abstract

Since the introduction of derivatives to the financial markets, volatility trading has emerged as a method for investors to make money in every market condition. Hedging methods have become crucial tools for liquidity providers operating in financial markets, emerging alongside the introduction of derivatives. Today, they play a parallel role alongside derivatives, ensuring the necessary risk mitigation and liquidity in the markets. The most commonly used hedging method is delta hedging which effectively eliminates the directional risk associated with options. Dispersion trading appears to be a profitable and precise method for hedging vega risk, offering an effective approach to manage and mitigate volatility-related concerns. The aim of this thesis is to investigate the effectiveness of dispersion trading in mitigating vega risk specifically in OMXS30 options. The approach involves implementing a backtesting strategy that encompasses taking a short position in OMXS30 index volatility while simultaneously adopting a long volatility stance on a tracking portfolio. This strategy ensures a net vega of zero. This investigation aims to determine if the dispersion trading strategy can be a reliable risk management tool. It was found that vega could accurately be hedged using dispersion trading. However, when considering the bid-ask spread, the strategy did not show profitability over the simulated period. Weighting the portfolio more in favour of companies with smaller bid-ask spreads did not show improved profitability.

Keywords

Dispersion Trading, Volatility Trading, Volatility Hedging,Vega Hedging, Option Trading, Back-testing, Liquidity provider, OMXS30 options, Index options

Sammanfattning

Sedan introduktionen av derivat på de finansiella marknaderna har volatilitetshandel dykt upp som en metod för investerare att tjäna pengar i alla marknadsförhållanden. Parallellt med introduktionen av derivat på de finansiella marknaderna har säkringsmetoder vuxit fram och är idag väsentliga instrument för de likviditetsgivare som är verksamma på marknaderna. Den vanligaste säkringsmetoden är delta säkring som tar bort den riktade risken i optionen. Att säkra vegarisken med spridningshandel tycks vara både en lönsam och pålitlig säkringsmetod. Detta examensarbete syftar till att undersöka effektiviteten av att använda en dispersionshandel för att minska vegarisken i OMXS30-optioner. Metoden involverar att simulera en strategi baserad på att vara kort volatilitet i OMXS30 och lång volatilitet på en spårningsportfölj på historisk data. Genom denna undersökning strävas det efter att avgöra om strategin för spridningshandel kan vara ett tillförlitligt verktyg för riskhantering. Det visade sig att vega kunde säkras med hjälp av spridningshandel. Strategin visade sig vara lönsam under den simulerade perioden men när köp- och sälj-spreadarna i de enskilda aktieoptioner inkluderades var det inte längre lönsamt att utföra metoden. Att vikta portföljen mer till förmån för företag med mindre köp- och sälj-spread visade inte på förbättrad lönsamhet.

Nyckelord

Spridningshandel, Volitilitetshandel, Volitilitetssäkring, Vega säkring, Optionshandel, Back-testing, Likviditetgivare, OMXS30 optioner, Index optioner

Acknowledgements

We want to express our sincere appreciation to all who have contributed to completing this Master's Thesis project. Firstly, we thank our advisor, Jan Sand, from KTH, for his valuable feedback, guidance, and support throughout our research. We also thank Henrik Hult for his unwavering commitment and invaluable contribution. We want to acknowledge SEB for providing us with the resources and opportunity to conduct our research. We are especially grateful to Patrik Lagerholm at SEB for his generous support and assistance throughout the process.

Stockholm, June 2023
Albin Irell Fridlund
Johanna Heberlein

Authors

Albin Irell Fridlund (albinif@kth.se) and Johanna Heberlein (jheb@kth.se)
Master's degree project in Financial Mathematics
KTH Royal Institute of Technology

Place for Project

Stockholm, Sweden
SEB

Examiner

Jan Sand
KTH Royal Institute of Technology

Supervisor

Jan Sand
KTH Royal Institute of Technology

Supervisor

Patrik Lagerholm
SEB

List of Figures

List of Tables

Contents

1 Introduction

This chapter introduces the aim of this study, including problem formulation, a brief background related to dispersion trading, research questions, purpose and goal where benefits, sustainability and ethics are covered. This section also includes the model process that builds on back-testing data provided by SEB, the stakeholders and delimitations of the project.

1.1 Problem formulation

Financial instruments such as derivatives are popular trading tools on the financial market since they offer investors more flexibility in risk exposure, leverage and tie up less capital. It is well known that there is a discrepancy between the implied volatility of index options and its subsequential realized volatility, which has been argued by numerous studies further discussed in section 2. One way for traders to profit from this difference is by selling options with repeated delta hedging. However, a more sophisticated method of taking advantage of this discrepancy is called dispersion trading. Dispersion trading is a strategy used to profit from the difference in volatility between two financial instruments. The strategy uses the fact that index options generally have higher implied volatility (IOIV) compared to the Markowitz implied volatility (MIV) for a portfolio of single stock options [28]. One approach would be to trade the dispersion between the implied volatility and the realized volatility of an option. Another approach is to take profit from the spread between the IOIV and the MIV of a tracking portfolio. The discrepancies between IOIV and MIV of their components frequently appear on large indices in the equity markets, including the S&P 500, DAX Index and OMXS30 [28]. The strategy has become common in the financial market for several reasons, some of them stated below:

- It allows investors to diversify their portfolios by placing positions on a large market index while protecting themselves from the risk of individual stocks.
- It gives investors a chance to profit from the difference between implied and realized volatility, which may be quite profitable if it decreases or

expands in their favour.

- It can be applied either by itself or in combination with other trading methods that make up a broader portfolio.

Given the complexity of the trading strategy, however, it is mainly used by institutional investors, hedge funds, and liquidity providers [13]. This thesis will investigate if vega risk (risk associated with a change in implied volatility) in OMXS30[1] options can profitably be hedged by trading volatility in a replicating portfolio. That is, a strategy will be conducted where the aim is to profit from the spread between the implied volatility of an index option and the implied volatility of its individual components while at the same time cancelling out the vega risk in the index option.

1.2 Background

The flexibility in risk exposure can be constructed using various options strategies, where multiple options (called legs) are used to create the exact payoff structure the investor seeks. Further, the risks in option positions are boiled down to five significant risk factors: delta, gamma, theta, vega, and rho. All of these are derivatives of the Black- Scholes formula [17]. When trading options through a liquidity provider, delta hedging is commonly used to protect the liquidity providers' directional risk. This is done by cancelling out the directional risk in the option by trading an equal amount of the underlying assets (usually stocks, ETFs[2] or futures contracts) in equal size and the opposite direction to the delta of the option [34]. The problem with this hedging method is that it only protects against the directional risk and does not offer any hedging against the change in implied volatility or the time value. Historical development in IOIV for OMXS30 compared to MIV of a tracking portfolio can be observed in Figure 1.1. As shown in Figure 1.1, implied volatility can change drastically, implying a need for hedging against this risk. Furthermore, a clear trend where IOIV trades above MIV of the tracking portfolio can be observed, implying that selling volatility on the OMXS30 index and buying volatility on the tracking portfolio is

[1]OMX Stockholm 30, is an index of the thirty most traded shares on the Stockholm Stock Exchange

[2]Exchange-traded funds (ETFs), is a fund traded on an exchange

profitable for this time period.

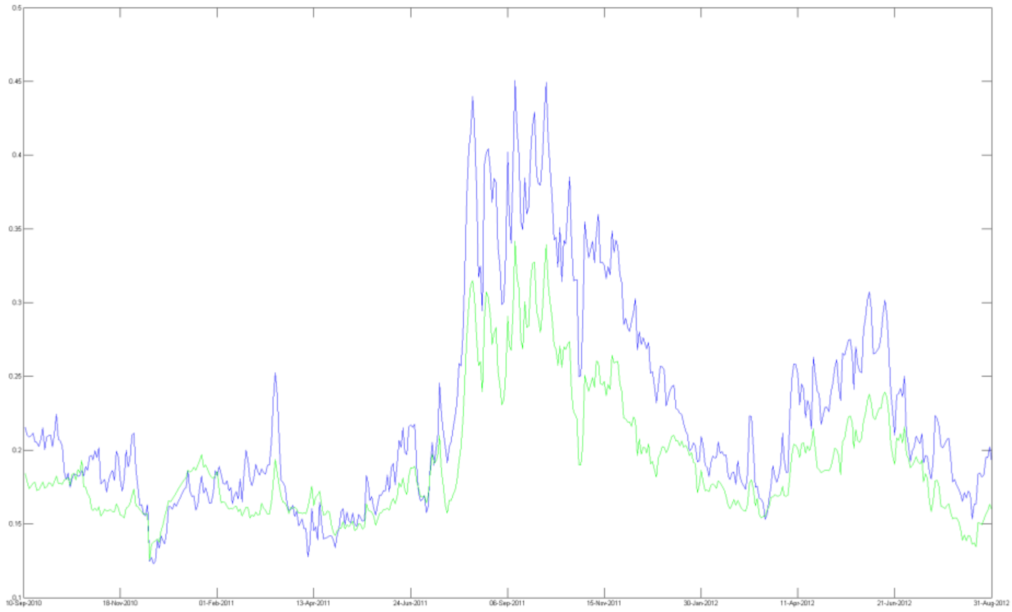

Figure 1.1: ATM Index options implied volatility (IOIV) in blue and Markowitz implied volatility (MIV) in green for OMXS30. Time period: Sep 2010 – Aug 2012 [24]

One way of hedging the implied volatility risk could be to trade options in the index's underlying components. This is referred to as a dispersion trade. Table 1.1 gives an overview of the different option types used in this model.

Table 1.1: Option types

	Single stock options	**Index options**
Underlying asset	Stock	Futures
Right to dividends	No	Discounted in future price

There are various ways to execute the strategy. However, one approach would be to buy a straddle (buy a call option and a put option) at the money (strike price is equal to or very close to the spot price of the underlying asset) with the same maturity on the index option, and to sell straddles on each component stock at-the-money (ATM) with the same maturity as the index options, and delta hedge the position.

1.3 Research questions

This thesis aims to investigate the following research questions:

1. How should a dispersion trade optimally be performed to hedge vega risks while minimizing cost and complexity?
2. How would the proposed strategy performe over historical data?
3. How does the strategy comply with the lack of liquidity in some single stock options?

1.4 Purpose

This thesis aims to investigate if there exist opportunities where traders can profitably hedge away vega risk in index options by trading in underlying components. This would reduce risk from volatility changes and improve SEBs' capability as a liquidity provider in index options.

1.5 Goal

The expected outcome from this research is a mathematical model, based on historical data, that figures out how a dispersion trade optimally would be set up to minimize the vega risk in index options.

1.5.1 Benefits, Sustainability and Ethics

This initiative has the potential to enhance the operational performance of SEB and facilitate the provision of superior prices for options to its clients. The firm can contribute to greater market liquidity by implementing more sophisticated hedging techniques and strengthening its market-making capabilities.

Profit-making strategies used in dispersion trading exploit market imperfections and price differences between particular securities and the market. As it could appear that money is made through market manipulation or by taking advantage of market situations, some may claim this method is unethical [16]. Dispersion trading could contribute to more trading activity in the underlying stocks, which can cause more market volatility and potentially harm investors not using this strategy, according to research by Huang and Zhang (2011).

1.6 Methodology

The methodology of this project focuses on back-testing the dispersion hedging strategy using historical financial data provided by SEB. The first step is to construct a tracking portfolio. It is decided to incorporate ten companies into the tracking portfolio, where the companies are selected using a forward selection method, with the variable R^2 being utilized in the selection process. The selected companies are included in multiple linear regression, and portfolio weights are calculated. The technique comprises purchasing straddles on the companies in the tracking portfolio and selling on the index. The method is based on taking position in- and holding options with six months to expiry and then rolling the cash position to new options with six-month expiry. Re-balancing to at-the-money (ATM) strikes is conducted when the strike price deviates by more than 5% from the spot price. The positions are re-weighted daily to achieve vega neutrality. To evaluate the method, the Greeks, implied volatility, and returns are calculated for each time step.

1.7 Stakeholders

The most significant stakeholders of this project are the shareholders, as a successful dispersion trading strategy could result in more significant earnings, higher dividends, and an increase in the value of the company's stock. Another significant stakeholder is the actors in the Swedish options markets. This group ranges from individual investors to hedge funds and pension funds. Better hedging methods would allow for better liquidity in the Swedish options market and thus allow market participants to lower their transaction costs when trading in these options. Lastly, the successful implementation of sophisticated hedging methods could give SEB the funds needed to further expand its corporate social responsibility initiatives, philanthropy, and environmental sustainability efforts. Thus, the local community is also to be treated as a relevant stakeholder.

1.8 Delimitations

When compiling a dispersion trading technique, it is essential to be aware of the restrictions and limits that apply to it. The study is nonetheless exposed to several restrictions and limitations. Due to the lack of data on option prices, SEB's dataset needed to be completed as missing data on bid-ask prices was discovered. During the investigation, it came to light that there were several instances of missing data. Days without data were solved by using the nearest available data point. The study is constrained by the availability of data provided, which is OMXS30 components between 2021-06-01 and 2023-03-24. However, because the simulation needs 90 days of stock returns from calculating the tracking portfolio, the simulation will be conducted over the period 2021-08-30 to 2023-03-24. As Atlas Copco B and Autoliv had no options, they were thus not considered in the model. Sinch was furthermore removed as well as the data provided was determined to be too limited. Further, the study is predicated on several models and assumptions that could not represent how financial markets behave. It is also presumed that American stock options may be modelled similarly to European stock options, with the underlying asset recalculated to the spot price of its forward contract. Commission-free transactions are also presumed.

1.9 Outline

A review of related and connected research is covered in Chapter 2 to provide the reader with a more comprehensive understanding of the topic and the state of the research. A theoretical background follows the literature review in Chapter 3, where relevant knowledge for this study is covered. Chapter 4, which pertains to the methodology section, comprehensively describes all methods and their respective implementation. The section also describes how the outcomes will be assessed and validated. Chapter 5 will present the investigation's results, while Chapter 6 will comprehensively discuss the findings. The study's conclusions and any possible directions for further investigations will be covered in the concluding chapter, Chapter 7.

2 Literature review

This literateur review will provide an overview of previous academic research on dispersion trading and how different approaches have previously been performed. The results from the methods are discussed and evaluated.

2.1 Strategies

Numerous studies have examined the best way to implement a dispersion strategy. FDAXHunter (2004) is one researcher who discusses different aspects of the strategy and makes simplifications to discuss the method further [13]. Magnusson (2013), Lisauskas (2011), Maze (2012) and Jokela (2022) investigate basic dispersion trading strategies, and Marshall B (2009), Deng (2008), Lisauskas (2011), Bakshi et al. (2003), Driessen et al. (2005) and Jokela (2022) are all researchers who disagree with the existence of an efficient market and argues that index options trade at higher implied and realized volatility compared to individual stock options. Fatemi & Krol (n.d.) and Choi, C. Y. (2008) are other researchers who examine the profitability of conducting vega and theta hedging within a dispersion trading strategy.

Ganatra (2004) focuses on implementing the strategy using variance swaps, which involves taking positions in options on individual stocks to take advantage of differences in volatility [14]. The study from Ganatra (2004) uses data from the Euro Stoxx 50 index[3] where it is confirmed that using variance swaps generates significant profits. However, after the financial crisis of 2008, the liquidity in single stock variance swaps has decreased, making the method irrelevant from today's standpoint [30]. Deng (2008) also investigated the strategy by the use of at-the-money (ATM) straddles on the S&P500[4] index where the implied correlation was employed as an indicator for when to initiate dispersion trades [9].

Lisauskas (2011) uses a technique where call options are traded on both the

[3]The performance of the top 50 firms among the 20 supersectors in terms of free-float market cap is reflected in the EURO STOXX 50 Index.

[4]The S&P 500 is a stock index of 500 large publicly traded companies trading in the United States.

index and components, and a delta hedge is then used to achieve delta neutrality [22]. Notably, the study agrees that using straddles or strangles, rather than just puts or calls, is more practical since they have a relatively minimal initial delta exposure due to their construction, which lessens the requirement for delta hedging. Magnusson (2013) performs a study evaluating three sorts of tactics; the straddle strategy, the strangle strategy, and a combination approach on the OMXS30 index. After examining the strategies, he found that the straddle had the most significant return and was the riskiest of the three trading techniques [24]. Nevertheless, all three strategies had positive returns, similar to the results reported by Lisauskas (2011). Call options and Cross-Sectional Volatility (CSV) swaps were used in the study conducted by Maze in 2012 on the South African options market. He discovered that the success of these trades is highly dependent on factors such as implied volatility, bid-ask spread, and trading costs. CSV swaps were found to be too volatile for practical use in dispersion trades [31]. Jokela (2022) evaluate both the correlation-based and volatility-based indicators to identify the best method. A discovery was made that those volatility-based indicators are the better choice when performing a dispersion trading strategy [20].

Furthermore, Fatemi & Krol (n.d) provided evidence that both vega-neutral and theta-neutral dispersion trading could be profitably implemented between 2008 and 2019 [12]. Vega neutrality was obtained from holding equal weights of vega notional of the index and total vega notional of the single stocks. This is also confirmed in a study by Choi, C. Y. (2008), where she discovered that performing both a vega-neutral and theta-neutral dispersion trading strategy was profitable [7].

2.2 Empirical edvidence

In 2008, Marshall (2008b) evaluated empirical data on dispersion trading and proved that Markowitz implied volatility (MIV) is lower than implied volatility on index options (IOIV). The research was performed on the S&P500 index and its underlying components for a two-year study period with data gathered from synthetic VIX indexes[5] provided by Bloomberg with at the end-of-day data. The study found that IOIV trades at higher levels than MIV. Marshall also revealed that increasing transaction costs entails decreased trading opportunities but that they were still existent [28]. Using VIX− indices, however, inconveniences traders as it is hard to apply when using different financial instruments since VIX− indices only measure the average volatility on a specific instrument over a 30−day period. Subsequently, traders cannot buy and sell all the options used to calculate the index. A dispersion trading strategy, therefore, involves individual stock options with different volatilities based on the current volatility smile and term structure, which is further discussed in section 3.11. To conclude, Marshall (2008b) proved that trading opportunities exist for a dispersion trading strategy. However, the study does not provide evidence for a practical implementation strategy or performance.

2.3 Mispricing

Although various researchers have supported the concept of an efficient market, mispricing is a phenomenon that economists and finance experts have extensively argued. Eugene Fama wrote an article in the 1970s where she assesses the empirical evidence for the concept and argues the impossibility of profiting from mispricings as the current stock price reflects all available information [11]. A critical overview of the efficient market is seen in an article by Malkiel (2003), where he claims that a market can be efficient in a short-term perspective but not in a long-term view [25]. Hence, he indicates that traders can gain deviant returns by identifying mispricing in the long run.

Bakshi et al. (2003) argue for one of the reasons for mispricing in the article "Stock return characteristics, skew laws and the differential pricing of individual equity options" [2]. He claims that mispricings are not fundamental flaws in the financial market but consequences of the complex process of investing. According to his theory, index options are priced higher than stock options due to the exposure to different risks, such as volatility and correlation risks. He argues that the index premium depends on the difference in risk-neutral

[5]VIX is a measure of the stock market's expectation of volatility based on S&P 500 index options

skewness between the underlying distributions of the options [2]. This is supported by Jokela, who made a study in 2022 on the Swedish market where he discovered that the OMXS30 index's implied volatility was higher than Markowitz's implied volatility [20]. The study by Lisauskas (2011), made on the DAX index, also confirms this hypothesis [22]. However, Maze (2012) found an unusual characteristic in the South African options market as he could not detect any appreciable difference between the implied volatility of index options and the implied volatility of the Markowitz equation [31].

Further, Driessen et al. (2005) argue that the risk premium for index options exists because index options can hedge correlation risk, which stock options cannot [10]. This, in turn, drives the risk premium for index options which makes them more expensive than stock options. Another hypothesis for mispricing is the inefficiency when pricing index and stock options, which in turn affects prices not to reflect the actual value of the underlying assets. According to a study by Deng (2008), if the price gap was caused by the pricing of some risk factors, as suggested by Bakshi & Kapadia and Driessen et al., there should not have been any difference in the profitability of dispersion trading techniques before and after the structural change [9]. Her study found that the monthly profit decreased by 24 percentage points after the 2000 market structure shifts during her research period of 1996 to 2005. This indicated that the inefficiency of the options markets was primarily to blame for the difference in pricing.

3 Theoretical Background

This chapter introduces relevant theory to the project to provide the reader with a comprehensive understanding of the theoretical framework that informs the research and to establish the context for the research question and objectives.

3.1 Dispersion trading

Dispersion trading is a form of a volatility-based trading strategy used on the financial market to seek profit from differences in implied volatility between financial instruments. The typical approach of this strategy involves selling volatility on an index while simultaneously taking opposing positions on a tracking portfolio linked to that same index. The goal is to profit from the possible discrepancy in implied volatility between the index and the its components [28]. The strategy becomes profitable due to the tendency of the implied volatility of index options (IOIV) to trade higher compared to the Markowitz implied volatility (MIV) associated with the tracking portfolio, as previously discussed by various researchers in section 2.

There are several ways to implement a dispersion trading strategy. One approach is to perform the strategy using variance swaps. This method was prevalent in the early 20s as it was highly profitable for brokers. However, following the financial crisis in 2008, investors switched to other approaches that would be more beneficial [30]. The most frequently used method today involves positions in options, as it allows traders to take leveraged positions on the dispersion spread. The strategy involves a short volatility position on the index and a long volatility position on the index components [9]. As trading and managing multi-leg option structures across all 30 components of OMXS30 would be difficult and incur significant transaction costs, this thesis will construct a tracking portfolio, including the most correlated stocks against the OMXS30 index. A dispersion trade is then constructed between the tracking portfolio and the OMXS30 index. Volatility will be traded using at-the-money (ATM) option straddles, where the number of straddles traded in the single stock options is determined by what cancels out the vega risk in the OMXS30.

Furthermore, implementations of dispersion trading can either have a constant exposure to the market (naive strategy) or a selective exposure based on when the MIV of the tracking portfolio is determined to trade at attractive levels [20]. Both variants have been shown to be profitable on historical data. The proposed strategy for this research will be of the naive type as it aims to provide constant vega protection.

According to the efficient-market hypothesis, the price of index options should be closely related to the price of its underlying asset as the value of the index is derived from the price of its constituent stocks [11]. Dispersion trading, however, argues against this theory as the strategy is based on mispricing due to differences in implied volatility between individual stocks within an index. As mentioned in section 2.2, Marshall (2008b) made a study on the matter and proved the deviation of implied volatility on the American S&P500 index [28]. Generally, index options trade at higher implied volatility than the corresponding portfolio of single stock options. This allows traders to profit from the dispersion between the two financial instruments, as index options are often more expensive than stock options.

Although financial research suggests that the strategy could be profitable, dispersion trading is a complex strategy that requires deep understanding before implementation. One significant risk that could occur while performing the strategy is if the market becomes heavily correlated, leading to difficulty gaining from the differences in implied volatility between the financial instruments. As changes in volatility and other factors on the market could significantly impact the profit versus loss, the strategy's risks can accumulate quickly if the trader is not disciplined and fails to monitor its positions [4]. Therefore, analysing the risk and potential rewards is essential before implementing a dispersion trading strategy.

3.2 Tracking Portfolio

A tracking portfolio, often referred to as a replication portfolio, is a portfolio of assets designed to closely match the performance of a benchmark index or a specific market. The first step when constructing a tracking portfolio is to choose what instruments to use and to allocate their respective weights. There are numerous techniques when constructing a tracking portfolio. However, the most common approach is through regression, where one can construct the portfolio as a linear or non-linear problem. Linear regression or principal component analysis (PCA) are two popular methods for modelling the tracking portfolio as a linear problem. In this thesis, the tracking portfolio will be constructed by linear regression, further discussed in section 3.6 where the tracking index is OMXS30, and the most correlated stocks will be used to replicate the portfolio.

Researchers like Krink et al. (2009) and Beasley et al. (2002) successfully used genetic algorithms to conduct their respective tracking portfolios in a non-linear approach. Krink et al. (2009) used differential evolution to build their tracking portfolio, whereas Beasley et al. (2002) developed evolutionary heuristics for their tracking portfolio [21] [3]. Although the non-linear technique requires more work since it frequently calls for more numerical solutions, it allows traders to include non-linear constraints, such as transaction costs, into their optimization model [26]. Different strategies have varied degrees of advantages and disadvantages. The critical difference is whether the problem is handled as a linear problem that can be solved simply or as a complicated model that requires more data, such as non-linear constraints. The extensive challenge of constructing a tracking portfolio has been well-argued by researchers. In a study by Blume and Edelen (2002), the reason for this occurrence is that index-tracking funds face the same issue as they represent significant proportions of managed assets [6]. Some academics contend that these issues could be resolved using a simple model while preserving the benefits of the cointegration process [33].

3.3 Correlation Matrix

When building a tracking portfolio, correlations in return between different components will appear. A covariance matrix describes the cross-correlations between the portfolio assets i, j and identifies the unique relationship between the coefficients [1]. As correlation coefficients can cause deviations from the index, it is essential to consider the correlations between the underlying assets to avoid problematic scenarios in the implementation and ensure that the tracking portfolio replicates the index and minimize the tracking error. One effect of highly correlated stocks is multicollinearity between the assets, which can lead to unstable estimates and affect the model's accuracy. As a result, correlation matrices are crucial because they allow investors to determine which portfolio holdings have a strong correlation to the index and which ones do not, which is necessary for a dispersion strategy to succeed.

$$\sigma^2 = \sum_{i=1}^{n} w_i^2 \sigma_i^2 + 2 \sum_{i=1}^{n} \sum_{i=1}^{n} w_i w_j \sigma_i \sigma_j \rho_{ij} \tag{1}$$

In Equation 1, the variance of a portfolio consisting of n assets is displayed, which was initially conducted by Markowitz in 1952 [27]. The weight of the i:th asset is represented by w_i, the variance of the i:th asset is described by σ_i^2 and ρ_{ij} is the covariance between stock i and stock j. The correlation matrix for the portfolio is then displayed as follows in Equation 2.

$$\begin{bmatrix} \rho_{i,i} & \rho_{i,j} & \cdots & \rho_{i,n} \\ \rho_{j,i} & \rho_{j,j} & \cdots & \rho_{j,n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n,i} & \rho_{n,j} & \cdots & \rho_{n,n} \end{bmatrix} \tag{2}$$

The variance equation's correlation matrix is symmetrical, and its coefficients, which range in value from -1 to 1, represent the degree and direction of the linear link between two assets. A correlation coefficient with a value of 1 denotes a perfect positive relationship where all of the assets move in the same direction and at the same rate. A correlation coefficient with a value of -1 denotes a relationship where all of the assets move in the opposite direction. If the coefficient has a value of 0, there exists no linear relationship [4]. As the diagonal represents the asset and its complete correlation, the diagonal of the matrix has a value of 1 at all times. Henceforth, in the case of portfolios containing a lot of assets, such as the OMXS30 or S&P500, the covariance matrix can quickly grow in size, which will cause much complexity. To address this problem, the correlation matrix can be scaled, and thereby, the complexity of large portfolios can be reduced. Standardization is one of the most commonly used methods when scaling a covariance matrix, and it involves dividing the mean by the standard deviation of each variable. This procedure guarantees that every variable is scaled equally and given the same weight during the process. As the correlation between asset i and j has the same relation as the correlation between asset j and i, the number of cross-correlations for a correlation matrix of n stocks can be decreased by making use of the correlation matrix's symmetrical nature and thus remove the unit diagonal points according to Equation 3.

$$\frac{(n \times n - n)}{2} \tag{3}$$

Further methods can be implemented to scale correlation matrices, and what decides the method depends on the purpose of the analysis [19]. According to Marshall (2008a), the risk of a portfolio can be divided into two categories; unsystematic risk and systematic risk [29]. The unsystematic risk is the risk exploited from one stock, which is the risk diversified when incorporated into a portfolio. In contrast, systematic risk is the risk that is shared by the market. With this approach, one can assume that the correlation coefficient between the tracking portfolio and the replicated index is close to 1. Thus, one can assume that the portfolio has minimal unsystematic risk and remove it from Markowitz's model and calculate the standard deviation according to Equation 4 where σ_m denotes the risk shared by the market.

$$\sigma_m = \sum_{i=1}^{n} w_i \sigma_i \rho_{i,m} \tag{4}$$

In this equation, the cross-correlations are reduced to n. From this point onward, it is only necessary to compute the correlation between the stocks in the tracking portfolio and the market portfolio, which represents the systematic risk.

3.4 Tracking Error

Tracking error is a measure used in finance to assess how closely an investment portfolio tracks its benchmark index. It is calculated as the standard deviation of the difference in returns between the portfolio and the benchmark over a specific time period shown in Equation 5. The portfolio's return is described by r_p, and r_b is the benchmark's return. By quantifying the deviation from the benchmark, the tracking error provides insights into the consistency and risk associated with the portfolio's performance.

$$\text{Tracking error} = \sqrt{Var(r_p - r_b)} \tag{5}$$

A higher tracking error indicates a greater divergence from the benchmark, suggesting increased risk. Contrarily, a lower tracking error signifies a more consistent replication of the benchmark [32]. To get the annualised tracking error, the tracking error is multiplied by the square root of the number of periods in a year, shown in Equation 6.

$$\text{Annualized tracking error} = \text{Tracking error} \cdot \sqrt{\text{Number of periods in a year}} \tag{6}$$

3.5 Liquidity Indicators

Liquidity indicators are often used in the financial market to rate the simplicity and effectiveness of trading and the marketability of various assets. A crucial market efficiency and stability component is liquidity, the degree to which an asset or security can be bought or sold on the market without materially changing its price.

3.5.1 Bid-Ask Spread

The bid-ask spread is one of the most commonly used indicators of liquidity. It measures the difference between the highest price a buyer is ready to pay for a security (the bid) and the lowest price a seller is willing to take (the ask). Since trading costs are generally minimal and buyers and sellers can transact at prices close to the asset's fair value, a narrow bid-ask spread typically implies that an asset is highly liquid [23]. Due to its often minimal bid-ask gap, currencies are regarded as the most liquid traded securities. Thus, for any stock or security, this criterion is frequently applied as a measure of liquidity. Empirical studies have been made on how the bid-ask spread should be measured, and the Demsetz model suggested a cross-sectional regression equation according to Equation 7.

$$s_i = \beta_0 + \beta_1 \ln(M_i) + \beta_2 (1/p_i) + \beta_3 \sigma_i + \beta_4 \ln(V_i) + \varepsilon_i, \tag{7}$$

Where s_i is the average bid-ask spread of the i:th security modelled as a function of variables: log market capitalization M_i, price inverse $1/p_i$, the riskiness of the security measured by the volatility of past returns p_i, and a proxy for activity such as log trading volume, V_i. Another concept of the bid-ask spread was modelled by Glosten and Milgrom (1985), who took the market makers' participation into account and suggested that there are two types of traders on the market, informed or uninformed traders [15]. The idea is that informed traders hope to profit from the uninformed traders' lack of market knowledge. While the market maker typically loses against knowledgeable traders, they make up for these losses on ”noise” trades. Trading for urgent financial demands, such as the need for cash flow or the desire to smooth out their consumption over time, is known as noise trading [23]. Assuming symmetry on the expected value of the bid p_t^{bid} and ask p_t^{ask}, the bid-ask spread will be represented according to Equation 8 where σ denote the range of uncertainty and ω indicate the consistent percentage of traders who have private information.

$$p_t^{ask} - p_t^{bid} = \omega\sigma \tag{8}$$

Other commonly used indicators include trading volume, which counts the total number of shares or contracts exchanged over a certain period. Since more buyers and sellers are actively participating in the market and orders can be executed promptly and effectively, higher trading volumes often indicate better market depth and liquidity.

3.6 Multiple Linear Regression

The fundamental concept behind a linear regression is to find an optimal linear combination between the dependent variable Y and the independent variables $X_1, X_2, \ldots, X_k$ with the lowest possible margin of error [19]. The dependent variable and independent variables are assumed to have a linear connection, and the residuals (the differences between the predicted and actual values) are assumed to be normally distributed with a constant variance and is modelled according to Equation 9.

$$Y_1 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \epsilon \tag{9}$$

The coefficients reflecting the slope of the line are represented by β_0, which is the expected value of Y when $X = 0$ and β_1, β_2,$\cdots$, and β_k are the coefficients that represent the average rise in Y that results from a one-unit increase in X. The error term ϵ serves as a catch-all for the defects of this simple structure, including the probability that the underlying relationship is nonlinear, the possibility of other factors influencing Y to change, and the likelihood of measurement error [19]. It is assumed that the error term is unrelated to X. As the regression coefficients β_0, β_1, β_2,$\cdots$,β_k are unknown, they must be estimated and are predicted according to Equation 10.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \cdots + \hat{\beta}_p x_p \tag{10}$$

$$\text{RSS} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{11}$$

The i:th residual, which is the difference between the i:th observed response value and the i:th response value predicted by the linear model in Equation 10, is therefore represented by the equation $e_i = y_i - \hat{y}_i$. The residual sum of squares (RSS) is defined according to Equation 11.

3.6.1 R^2 statistics

In regression, R^2 is a statistical measurement independent of Y, used to assess how well a regression model fits the data; a higher R^2 value denotes a better fit [19]. It takes values between 0 and 1 and is defined according to Equation 12,

$$R^2 = \frac{\text{TSS} - \text{RSS}}{\text{TSS}} = 1 - \frac{\text{RSS}}{\text{TSS}} \tag{12}$$

Where $\text{TSS} = \sum (y_i - \bar{y})^2$ is the total sum of squares and RSS is described in Equation 11. The R^2 measure can also be adjusted, so-called adjusted R^2. This statistical measure is used to penalize large amounts of regressors and is defined according to Equation 13.

$$R^2_{Adj} = 1 - \frac{\text{RSS}/(n-d-1)}{\text{TSS}/(n-1)} \tag{13}$$

Where n is the number of data points, and d is the number of companies included in the model. Compared to the ordinary R^2 value, the R^2_{Adj} value is always lower. However, they work in similar ways as a large value of R^2_{Adj} indicates a model with small test error in opposition to BIC, AIC and C_p, which are other examples of statistical measurements where a large value implies a large test error [19].

3.6.2 MSE and RSME

Mean Squared Error (MSE) measures the average squared deviation between the regression line and the predicted values. MSE is thus a measure of the prediction error in the model and is calculated as can be seen in equation Equation 14.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{f}(x_i)\right)^2 \tag{14}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{f}(x_i)\right)^2} \tag{15}$$

Where the prediction made by $\hat{f}$ that gives for the ith observation is denoted by $\hat{f}(x_i)$. Taking the square root of the MSE results in Root Mean Squared Error (RMSE). This figure is more intuitive to compare against the predicted value as it is not squared [19]. The calculation for RSME can be seen in equation Equation 15.

3.7 Black-Scholes Model

The Black-Scholes model, also known as the Black-Scholes-Merton model, is a mathematical formula used in the financial market to price derivatives such as option contracts. The model is a linear parabolic partial differential equation (PDE) and is demonstrated in Equation 16 [5].

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0, \tag{16}$$

The theoretical price $V(t, s)$ depends on five input variables; the underlying stock price of an asset S_t, the volatility σ of the price, the expiration time t and the risk-free interest rate r. It is assumed that the underlying stock price S_t follows a geometric Brownian motion with constant volatility σ, i.e. the returns are normally distributed with constant volatility over time [27]. The derived model is also assumed to have no transaction costs, no dividends paid by the underlying stock and that the risk-free interest rate is constant over the life of the option [1]. Formulas for the costs of European call and put options are perhaps the most well-known solutions to the Black-Scholes model and take form according to Equation 17 and Equation 18 with maturity time T.

$$c = S_0 N(d_1) - Ke^{-rT} N(d_2) \tag{17}$$

$$p = Ke^{-rT} N(-d_2) - S_0 N(-d_1) \tag{18}$$

The letters c and p denote the European call and European put price, $N(x)$ is the cumulative probability distribution function for a variable with a standard normal distribution, K is the strike price, d_1 and d_2 are formed according to Equation 19 and Equation 20 [17].

$$d_1 = \frac{\ln(S_0/K) + (r + \sigma^2/2)\,T}{\sigma\sqrt{T}} \tag{19}$$

$$d_2 = \frac{\ln(S_0/K) + (r - \sigma^2/2)\,T}{\sigma\sqrt{T}} \tag{20}$$

The put-call-parity demonstrates the relationship between the European call and put stated in Equation 21 if the two options have the same time to maturity T and strike price K.

$$c + Ke^{-rT} = p + S_0 \tag{21}$$

3.8 Greeks

The risks associated with an option are generally expressed as the Greek symbol of delta δ , gamma γ , vega ν, theta θ and rho ρ . Each Greek symbol shows how the option will be affected by changes in the different factors determining the value of an option according to the Black-Scholes formula [17].

3.8.1 Delta

Delta shows how the option relates to the underlying asset's directional risk. Delta tells the investor how the options price will change if the underlying asset moves with one currency unit. A delta 0.5 option would thus increase in value to USD 0.5 if the underlying asset increased to USD 1 [17]. Delta can be negative and positive, ranging from 0 to 1 for call options and -1 to 0 for put options.

3.8.2 Gamma

Gamma measures the delta changes given one currency-unit change in the underlying asset [17]. Gamma is typically higher when the underlying trades around the option strike price (the option is at the money) and lower when the underlying trades at levels further away from the strike price. Gamma is also higher for options closer to expiration than for options further to expiration.

3.8.3 Vega

Vega measures how the option price is affected by a 1%-unit change in implied volatility. The price impact from the vega can be calculated by multiplying the change in implied volatility with the value of vega [17]. Adding the result to the option price gives the value after the change in implied volatility. Vega is most significant when the option is at the money and declines as the underlying trades further away from the strike price. Vega also declines as the time to maturity decreases.

3.8.4 Theta

Theta is a measure of how the options price relates to time. The value of theta represents the change in value after one day's worth of time [17]. Theta is generally negative and decreases in value more rapidly as the option matures. Theta is generally higher for at-the-money (ATM) options.

3.8.5 Rho

Rho serves as a measure of sensitivity, indicating how the price of an option is influenced by a 1% change in the risk-free interest rate. For long positions in calls, Rho is positive and tends to rise as the stock price increases. Conversely, for long positions in puts, Rho is negative and approaches zero as the stock price rises [17]. Typically, Rho holds greater significance (more positive for calls and more negative for puts) for options with longer expiration dates.

3.9 Forward pricing

A forward contract is a two-way agreement that calls for the delivery of the underlying commodity at a later time and at the agreed-upon price, with payment of the cash due at or after a predetermined number of days after delivery. Forward contracts are typically traded over the counter between two financial institutions or between an institution and one of its clients [17]. Figure 3.1 visualises payoff from a forward contract.

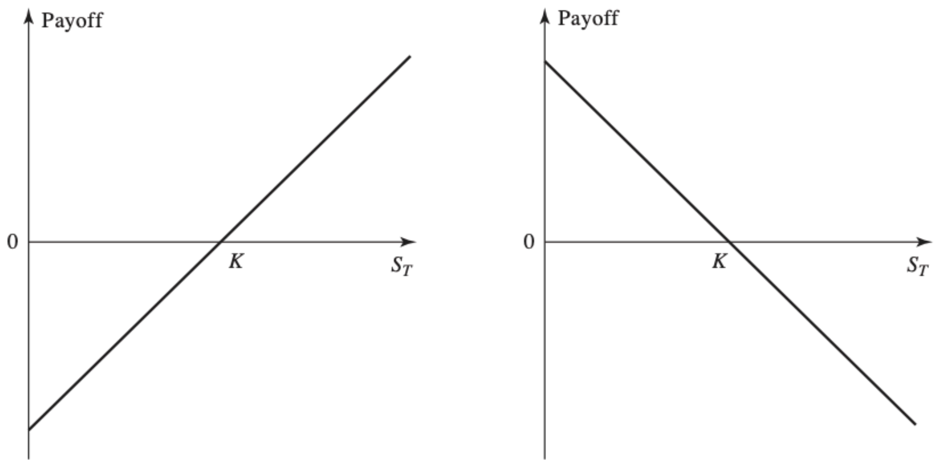

Figure 3.1: Payoffs from a forward contract; long position to the left, short position to the right [17]

By securing the price that the hedger will pay or receive for the underlying asset, forward contracts aim to eliminate risk. Contrarily, with the agreed-upon price, forward contracts also involve risk as future negative price swings may occur. For forward rates with known income, such as stocks paying known dividends, a general equation will take form according to Equation 22.

$$F_0 = (S_0 - I)\, e^{rT} \tag{22}$$

When valuing a forward contract, several parameters are included, such as time to delivery date in years T, the price of the underlying asset S_0 today, the price of today's forward contract F_0 and the risk-free rate r of interest per annum, calculated using continuous compounding, for security maturing on the delivery date. An expression of the value of a long forward contract on an asset with known income with present value I is represented in Equation 23.

$$f = S_0 - I - Ke^{-rT} \tag{23}$$

Where I is the present value of the known cash income, an arbitrageur can make a profit if $F_0 > (S_0 - I)\, e^{rT}$, by purchasing the asset and selling a forward contract on the asset or if $F_0 < (S_0 - I)\, e^{rT}$ the investor can profit by selling the asset and take a long position in a forward contract [17].

3.10 Volatility

The statistical measure of annualized spread within a stochastic process that models logarithmic returns of assets is called volatility. The most frequently used metric for quantifying the degree of a spread around the mean in a distribution is the standard deviation, a measure highly related to volatility. The two concepts can be associated with the square root of time rule; the convention between the two instruments, however, only applies if the returns are i.i.d[6] [1]. In finance, volatility is used to reflect the degree of uncertainty or risk associated with the price movement of an underlying asset over a certain period [1]. A high degree of volatility denotes that the security valuation can exhibit substantial oscillations over a brief time frame. In contrast, a diminished degree of volatility suggests that the value of the financial instrument remains relatively stable.

3.10.1 Implied Volatility on index options

Implied volatility is an essential concept in options trading as it refers to the market's expectation of future volatility of the underlying asset based on the current price of its options. As index options are based on the option of a portfolio rather than the asset itself, several studies have shown that the implied volatility of index options can provide essential insights into future movement in the overall stock-market [17]. Given that index options are of the European type, one does not need to consider early exercise or Asian tails. Hence, the most used method when calculating the implied volatility of index options is the Black-Scholes method described in section 3.7. Since it is impossible to solve the differential equation directly for the volatility of the underlying instruments, an iterative method is utilized. The Black-Scholes model calculates the resulting price based on the initial volatility that traders assume. This price is then

[6] A set of random variables is independent and identically distributed in probability theory and statistics if each random variable has the same probability distribution as the others and they are all independent of one another [17].

compared with the current market price, and based on the comparison results, the process is iterated with a more accurate implied volatility estimate [17]. One can choose more complex methods, such as the finite difference method, binomial trees or Monte Carlo, to estimate the implied volatility, but the iterative process is more effective numerically. In this project, the implied volatility will be estimated through the Newton Rapson method.

Moreover, the Black-Sholes model assumes no dividends are paid on the underlying instruments [17]. As OMXS30 options have delivery in the form of futures contracts, this assumption is met from the beginning, as futures contracts do not give the investor the right to dividends. Despite this, the value of the underlying used in this simulation will be the calculated forward value received from calculating the index's adjusted future, according to Hull (2018), with expiration at the same date as the option [17]. This has to do with the fact that the single stock options will need to be re-calculated this way, and doing the same for the OMXS30 index will give more comparable results.

When valuing options, another important consideration is the risk-free rate which is determined by various factors such as the yield curve. The yield curve visually represents the term structure of interest rates that varies with the time to maturity [8]. To simplify, it is a dynamic concept that represents the change in yield of bonds issued by the same issuer depending on the time to maturity. Although the yield curve can show fluctuations, it still seems to follow a specific pattern as rates with longer maturity times are higher than rates with shorter time to maturity and are of great importance when valuing options [17].

3.10.2 Implied Volatility of single stock options

When determining the value of European stock options, the Black-Scholes formula can be utilised to calculate the implied volatility. However, single stock options in Sweden are generally of the American option style, with no analytical evaluation approach. Therefore, numerical solutions or approximations are necessary for these types of options. American-style options differ from European stock options as European options provide the holder with the right, but not the obligation, to buy or sell the underlying asset at maturity. In contrast, American options allow early exercise and can be exercised at any time prior to its expiration date [17]. Exercising an option before its expiration date is generally not advantageous if no dividends are involved. In detail, American call options will always be more beneficial to keep until their expiration date as the investor will earn the risk-free rate. However, with American put options, early exercise might be more beneficial since waiting until the option expires to sell runs the risk of costing the investor the risk-free rate. Subsequently, American options have the drawback that, in the event of discrete payouts, early exercise may be advantageous if the dividend is sizable enough and the ex-dividend date is close to the maturity date. As the valuation of these types of options becomes too complex to solve analytically as it requires constraints [1], numerical approaches are used to solve these options. The three most common ways to value American options are:

- Binomial trees
- Monte Carlo simulations
- Finite difference method

In this thesis, American-style single stock options will be assumed to be modelled as European options accurately. The adjustment made to make this simplification more believable is to re-calculate the underlying asset's price to the price of a forward contract with the same expiry as the option. This will remove the issue with discrete dividends in the simulation.

3.11 Volatility smile and term structure

The implied volatility of options, which represents the market's prediction of the future volatility of the underlying asset, exhibits a well-known pattern known as the volatility smile. The volatility smile states that a European call option shares the same implied volatility as a European put option if the options have the same time to maturity [17].

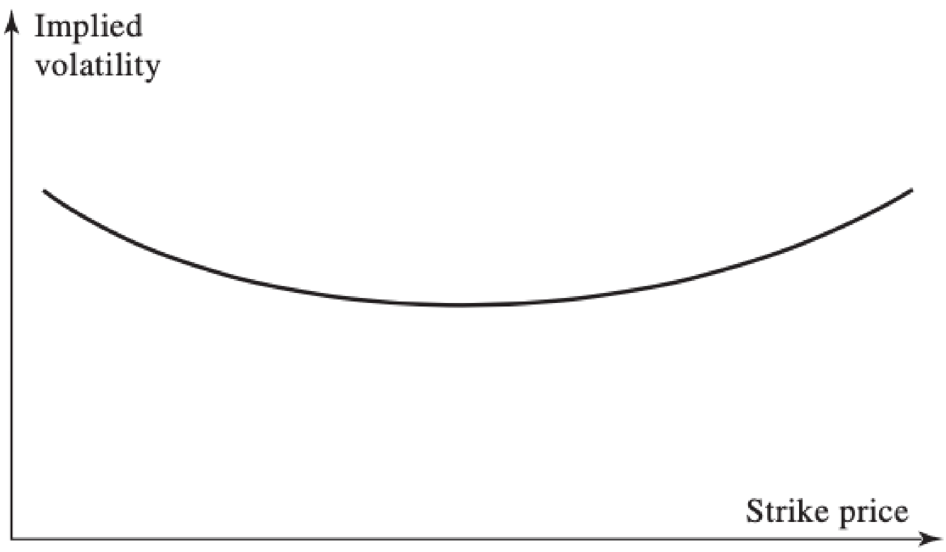

Figure 3.2: Volatility smile [17]

According to the Black-Scholes model, underlying security has a log-normal distribution with constant volatility, and for options which are at-the-money (ATM), the implied volatility is relatively low. However, as the options move further out-of-the-money (OTM), the implied volatility tends to increase, which results in a ”smile” if the volatility curve is symmetrically displayed in Figure 3.2 or in a ”skew” pattern if not symmetrical. This implies that the Black-Scholes model might not correctly value options, particularly those that are into the money or out of the money. As a result, when evaluating options and mitigating their risks, options traders must consider the volatility smile.

3.11.1 Volatility Term structure

Traders let the implied volatility be influenced by both the strike price and the time to maturity [17]. This relationship is demonstrated by the volatility term structure where options are organized according to their time to maturity with their respective implied volatility on the y-axis. An example of this is demonstrated in Figure 3.3.

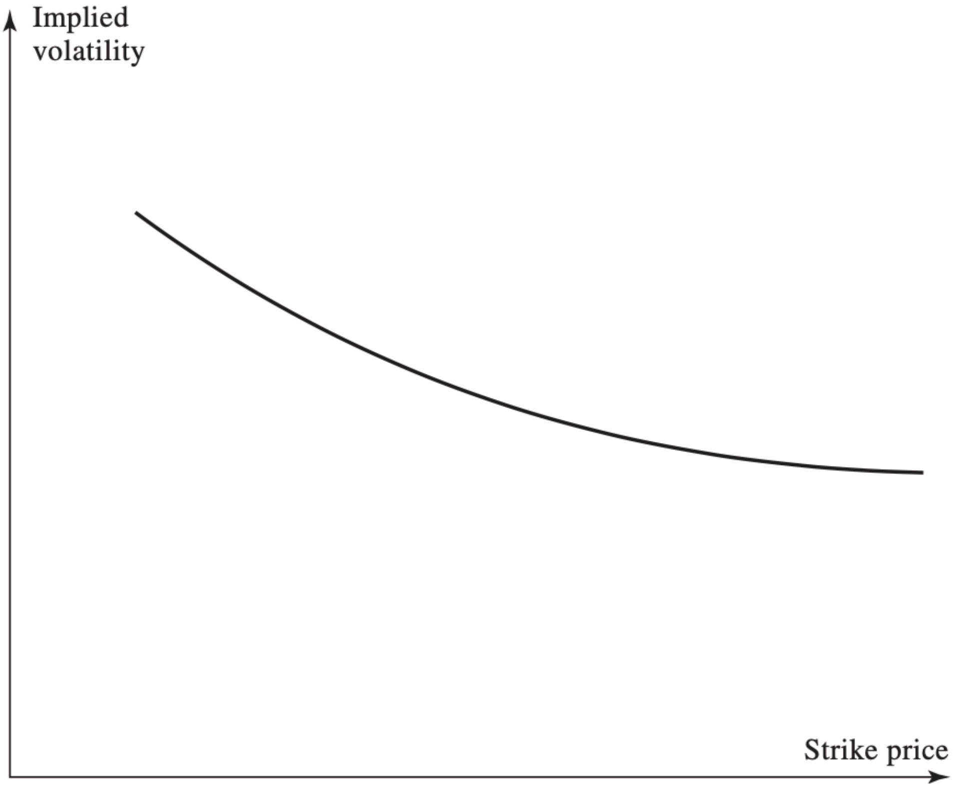

Figure 3.3: Volatility term structure for equities [17]

When short-dated volatilities are historically low, empirical data reveals that implied volatility is positively associated with time to maturity. This is because volatility is anticipated to rise in the following years. In contrast, implied volatility is inversely linked with time to maturity when short-dated volatility is historically high because there is a belief that volatility would decline with time [17]. Therefore, when pricing options with various maturities, traders consider both the current volatility level and projections for future changes.

3.12 Strategies

3.12.1 The straddle

A straddle is an options strategy that involves the simultaneous purchase or sale of a call and a put option with the same maturity, strike price, and underlying asset. Figure 3.4 depicts the resulting payoff diagram. This strategy is commonly employed by investors who anticipate a significant movement in the underlying asset's price in either direction, exceeding the cost of the premium paid for the strategy [17]. In essence, the straddle can be viewed as a wager on the trading range of the underlying asset until maturity.

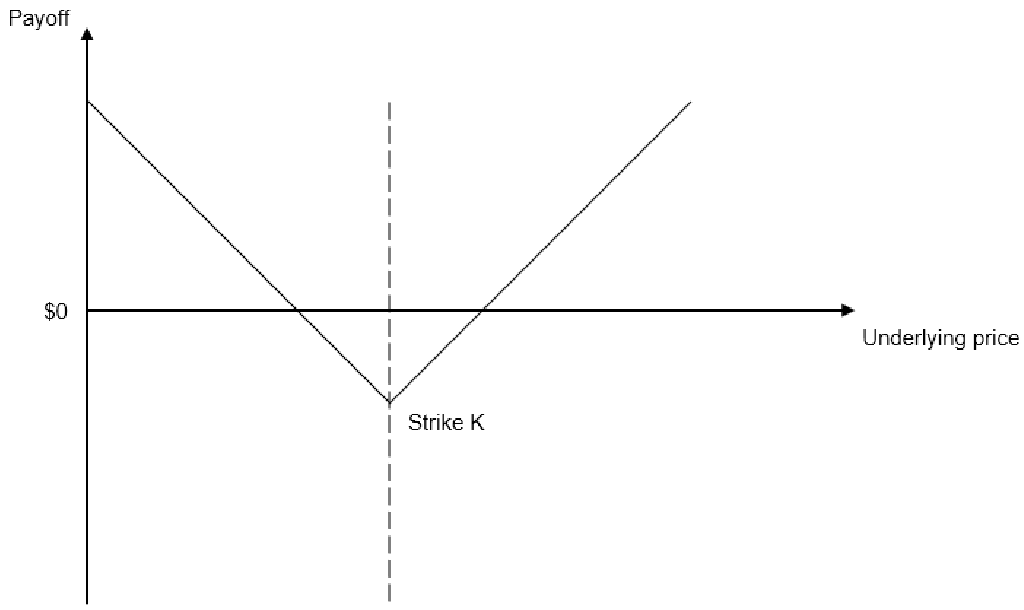

Figure 3.4: Straddle payoff

3.12.2 The strangle

The strangle is a financial strategy that shares similarities with the straddle, as both involve the purchase or sale of a put and a call with the same maturity. However, a key difference between the two is that in a strangle, the put option has a lower strike price than the call option. The resulting payoff graph is depicted in Figure 3.5. The strangle strategy may be employed by investors who anticipate a substantial movement in the underlying asset in either direction.

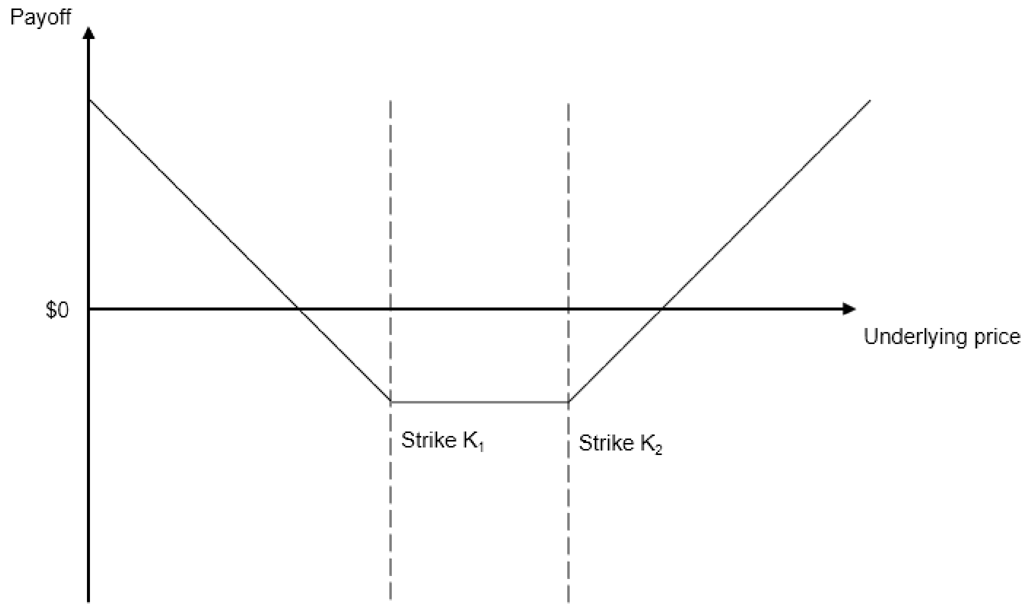

Figure 3.5: Strangle payoff

It is important to note that the strangle typically requires a larger movement in the underlying asset compared to a straddle in order to generate a profit [17]. Nevertheless, the premium paid for the strangle is typically lower, reflecting the increased difficulty of achieving a profit due to the greater movement required.

3.12.3 Delta hedging

Delta hedging is a widely used risk management strategy in financial markets to reduce or eliminate the directional risk associated with an option position. The technique involves taking a position in the underlying asset that is opposite in direction to the option, in an amount equivalent to the delta of the option [17]. The delta refers to the sensitivity of the option's price to changes in the underlying asset's price. To achieve delta neutrality, the position in the underlying asset is calculated as in Equation 24.

$$-(\delta \cdot k \cdot n). \tag{24}$$

Where δ is the options delta, k is the multiplier in the option, and n is the number of options contracts traded. It is worth noting that delta hedging must be adjusted regularly to account for changes in the option's delta caused by gamma. By employing delta hedging, investors can limit their exposure to changes in the underlying asset's price and reduce the overall risk associated with their option position.

4 Method

The general methodology is based on a backtesting of the dispersion hedging method on historical financial data. The following section will describe each step of the construction of the hedging method and methodology for conducting the backtest.

4.1 An overview

According to the literature review in section 2, it is proven to be more profitable to hold a short position on index volatility and a long volatility position on the single stock components rather than the other way around. The strategy is therefore expected to profit from the dispersion in implied volatility between single stock options and index options. This is possible by selling volatility expensive on the index options and hedging the risk away cheaper on the replicating portfolio, i.e. straddles are sold on the index and bought on each component. This study involves an initial short position of 100 SEK worth of OMX30 straddles where the value of 100 SEK in the index option is kept constant throughout the simulation. The number of single stock options to be sold will then be determined on how much would be needed to cancel out the vega risk in the 100 SEK position in OMXS30. For comparison, a similar strategy will be simulated with a fixed amount of 100 SEK allocated to both the index options and the single stock options. As the strategy is intended to be used as a hedging method against short vega in the OMXS30 index, a naive strategy is proposed, as it provides continuous hedging. A naive strategy in this context means that constant exposure of short volatility on the OMXS30 and long volatility on the tracking portfolio will be held to expiry and then repeated on the following six-month expiry options.

Multiple linear regression is used together with forward selection based on R^2 to select the ten companies with the most significant explanation power of the variation in the OMXS30 index. Multiple linear regression is chosen over alternative methods like lasso regression and PCA as it results in a more stable portfolio over time, and the number of companies included can be fixed, meaning less complexity. Further, using multiple linear regression, a replicating portfolio is constructed based on the 90-day lagging correlation between the OMX30 index returns and the returns of its components. An assumption for the simulations is that no commission will be paid. However, the strategy will be simulated, given that the full bid-ask spread is paid. Including the spread in the simulation is crucial due to its significance as one of the primary costs associated with options trading [20]. Commission costs can vary between contracts, and how much institutional traders pay is generally not disclosed, making it hard to include them in the simulation accurately.

The coefficients provided by multiple linear regression are re-calculated to the percentage weights of each company. The weights are then re-weighted to companies with smaller-than-average bid-ask spreads. Then lastly, the amount of money invested in option straddles on the replicating portfolio is determined by what would cancel out the vega risk in the index position. For comparison, the vega-weighted dispersion strategy will be paired with a strategy based on holding an fixed position of short 100 SEK on the index and long 100 SEK on the components. Further, options are traded based on the weights determined in the previous step. This thesis investigates six-month maturity options since more extended maturity options imply higher vega risk [17]. During each iteration of the simulation, the single stock options and index options are, if needed, re-balanced based on three different terms:

- Expiration: The options are re-balanced to a new option with a six month expiry if the current position expires the following trading day.
- Moneyness: The options are re-balanced to ATM strikes if they are more than 5% in the money, or more than 5% out of the money.
- Weight: For index options, positions are adjusted to ensure a 100 SEK exposure, while for single stock options, positions are adjusted based on replicating portfolio weights that cancels vega risk.

To investigate the affects of option liquidity, the bid-ask spread for each option in the tracking portfolio is calculated and compared to the average bid-ask spread of the portfolio. A penalty will then be introduced and simulated for companies with a more extensive bid-ask spread than the average. The implied volatility is approximated for each timestep by applying the Black and Scholes formula to the option's current market value. The Newton-Raphson method is then utilized to solve for implied volatility as it is effective in finding roots of non-linear equations. If the roots are not found, realized volatility will be used instead.

Lastly, delta hedging removes the delta risk and gives the results a more accurate representation of the effect received from hedging vega according to the proposed method. An overview of all the steps can be observed in Figure 4.1, where a more detailed explanation for each step in the strategy is found in the sections below. The profitability will be evaluated based on the portfolio returns over the simulated period, where the portfolio returns will be calculated based on two simulations: (1) it is assumed that the trader pays the entire bid-ask spread in each transaction, (2) it is assumed that the trader does all the transactions in the middle of the bid-ask spread.

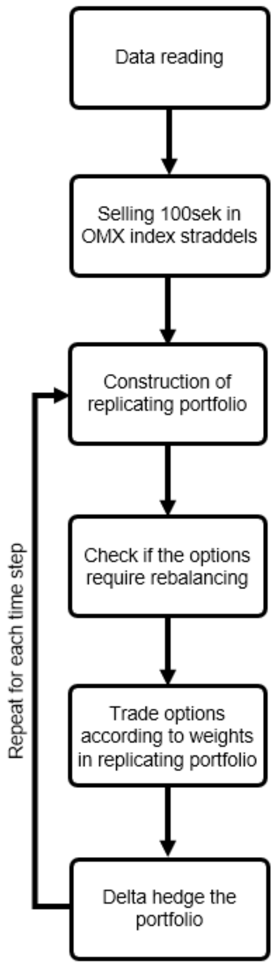

Figure 4.1: Overview of methodology

4.2 Data

The data utilised in this project is provided by SEB. It contains daily closing prices for the OMXS30 index and each associated component stock, the last quoted bid-ask spread for the OMXS30 index and each related option except for Atlas Copco B and Autoliv, SEBs risk-free rate and dividend estimated for each index component, the spot exchange rate for USD, Euro and Swiss crowns to SEK. The data contains closing prices from 2021-06-01 to 2023-03-24. A complete set of companies included in OMXS30 over the selected period and whether they have options can be seen in Appendix A. Furthermore, Sinch has been excluded from the simulation due to limited data.

4.2.1 Missing data

Removed data: The data over option prices contained several data points with the bid or ask prices named ”nan”. These data points are removed as they provide no value for the simulation.

Simulated data: The data sets used in this study do not include a complete set of bid-ask prices for all publicly traded term structures of the options that are being investigated. To receive prices where no data is found, the implied volatility from the bid and ask side of the four-term structures with the closest strike price are used to fit a second-degree polynomial. The decision to use a polynomial to approximate the unknown term structure is made because the volatility smile constitutes a second-order polynomial shape. Furthermore, in-the-money (ITM) calls and out-of-the-money (OTM) puts share the same implied volatility curve, and ITM puts, and OTM calls share the same implied volatility curve. OTM calls and ITM puts are approximated separately from ITM calls and OTM puts to minimise errors in the approximation from possible skewness in the volatility curve. The implied volatility is calculated by solving the Black and Scholes formula iteratively using the Newton-Raphson method. The simulated option price is then received from the following steps:

1. The implied volatility is calculated for the four selected term structures. These volatilities are then used in a cubic spline interpolation to approximate the implied volatility of the term structure without price data. The implied volatility for the bid and ask sides are estimated separately to simultaneously approximate the spread in the term structure without price data. This was achieved using the cubic spline function in the SciPY library. Cubic spline interpolation was chosen over linear approximation methods, as it is known that the volatility curve has the shape of a second-degree polynomial. A linear approximation method would thus involve a larger error term than what can be achieved from cubic spline interpolation.

2. The bid and ask prices are then calculated using the Black and Scholes equation for the sought-term structure.

4.3 Dividend management

Since the most relevant option theory is based on the Black and Scholes formula, a simplification is made where the American-style single stock options are assumed to be represented as European-style options. This assumption is necessary since single stock options have discrete dividends, and Black and Scholes assume that no dividends occur during the period it covers. Since dividends will take place over the simulation, the dividends are discounted in the forward price and assumed to be continuous over the calculated period. The forward prices are also calculated on the OMXS30 index to minimise the risk of price differences. SEB's risk-free rate is used in forward pricing, and the spot prices for each exchange rate are assumed to be a valid approximation when converting foreign dividends to SEK.

4.4 Tracking portfolio

As trading every component in the OMXS30 index would infer unnecessary transaction costs, create increased complexity and potentially mean that each component might not be liquid enough for trading a more significant amount of options, a tracking portfolio is constructed to replicate the movements in the OMXS30 index. The selection process of companies to be included in the tracking portfolio is done by performing a forward selection of the ten companies, which is the best method to increase the R^2 value of the model the most. The portfolio is decided to be fixed at ten companies as it is thought to keep an adequate correlation factor throughout the simulation, and sizing up and down the portfolio infer unnecessary transaction costs. Further, the selected variables are included in a multiple linear regression to predict the return of the OMXS30 index. Forward selection based on R^2 together with multiple linear regression for creating the tracking portfolio has successfully been implemented by Magnusson [24] and proved to result in a stable tracking portfolio over time.

Each company's correlation coefficient β_j is received and re-calculated from the multiple linear regression to decide the number of contracts that should be traded in each name. Greater allocation is given to companies with better liquidity by penalising companies with larger bid-ask spreads than the average in the tracking portfolio with a constant k. This is done according to Equation 25 to 29 where Adj_i in Equation 25 is the adjustment factor for component i that allocate greater weights to companies with smaller bid-ask spread and smaller weights to companies with larger bid-ask spreads than the average. The bid-ask spread s_j is represented by a fraction shown in Equation 26.

$$Adj_i = \frac{1}{\left(1 + \left(s_i - \frac{1}{n}\sum_{i=1}^{n} s_j\right) \cdot k\right)} \tag{25}$$

$$s_j = \left(Ask_j - Bid_j\right)/Ask_j \tag{26}$$

To calculate the adjusted weights in percentage for each component received from the multiple linear regression, the adjustment factor in Equation 25 is multiplied with the percentage distribution, where B_i is the distribution for component i. This is done according to Equation 27.

$$G_i = Adj_i \cdot \frac{B_i}{\sum_{j=1}^{n} B_j} \tag{27}$$

To adjust so that the percentage distribution has a sum of 100% the adjusted percentage distribution G_i for each component i is divided to the sum of all adjusted percentage distributions according to Equation 28.

$$A_i = \frac{G_i}{\sum_{j=1}^{n} G_j} \tag{28}$$

In this simulation, an arbitrary amount of 100 SEK will be shorted on index straddles and the amount bought on the single stock options is determined by what removes the vega risk on the index options. To get the allocation in SEK, the percentage allocation A_i is first divided by the current mid price of the call and put option bought from a previous iteration, or if it is the first iteration in the simulation, a six-month at-the-money (ATM) put, or call is used. Furthermore, Equation 29 is multiplied by a constant to determine the number of contracts per company. This constant symbolises the total investment in buying straddles on the replicating portfolio. The optimal value of the constant is determined by solving the root of what total investment minimises the vega risk. To receive this, the net vega is calculated for the index options and compared to the net vega exposure in the replicating portfolio.

$$C_i = \frac{A_i}{mid\,price\,call_i + mid\,price\,put_i} \cdot Total\ Investment \tag{29}$$

The weights of the tracking portfolio are recalculated daily in the backtest.

4.5 Re-balancing and rollover

As the strategy is based on holding ATM straddles, re-balancing is done based on the last quoted price daily. The options position is taken on the first day of the backtest, and if the option stays within 95-105% moneyness, hold until the penultimate trading day of the option's maturity. On the penultimate trading day, the position is rolled over to new options with six months until maturity. The maturity of the traded option is six months, where the reason for choosing long maturity options is because they are more affected by changes in vega [17]. When rolling the options to new maturities, re-balancing to ATM strikes and correct weights are done simultaneously. The maturity date will stay the same if only re-balancing to ATM strikes are necessary, but the position is re-balanced to the correct weight in the same step.

4.6 Option Greeks

To evaluate the performance of the strategy option, Greeks on a portfolio level are calculated daily. In the calculation of the Greeks, a simplification is made where the single stock options, which are of American style, are estimated to be accurately modelled as European options. The calculated spot price of the associated forward contract replaces the spot price of the stock/index. The reason for using forward pricing instead of the spot price of the stock/index has to do with the fact that the Black-Scholes formula cannot handle discrete dividends.

4.7 The dispersion strategy

A summary of each step in the complete strategy used in this project can be observed below;

1. Daily returns from the last 90 days are calculated for the OMXS30 index and each stock component.
2. Straddles on the OMX index are being sold at a fixed position of 100 SEK.
3. Forward selection based on R^2 is used to select the ten companies with the highest correlation to the OMXS30 index.
4. The returns from the selected companies are regressed against the returns of the OMXS30 index using multiple linear regression.
5. The weighting of each component is based on the coefficients obtained from the multiple linear regression.
6. At each time step, the components in the tracking portfolio are examined to determine whether they still exhibit the highest correlations and whether they should continue to be weighted accordingly.
7. As the tracking portfolio should consider option liquidity, the bid-ask spread on each component option is calculated and compared against the average bid-ask spread in the tracking portfolio.
8. In each time step, the options held are assessed based on Expiration,

Moneyness and Weight.

- Expiration: The options are re-balanced to a new option with a six-month expiry if the current position expires the following trading day.
- Moneyness: The options are re-balanced to ATM strikes if they are more than 5% in the money or more than 5% out of the money.
- Weight: For index options, positions are adjusted to ensure a 100 SEK exposure, while for single stock options, positions are adjusted based on replicating portfolio weights that cancel the vega risk.

9. Implied volatility is estimated in each time step by solving the Black and Scholes formula given the current market value of the option. The Newton-Raphson method is used when solving for implied volatility.

5 Results

This section presents the results from the back-test of the dispersion trading method where straddles are sold on OMXS30 index and bought on the tracking portfolio. The results include profitability, greeks, implied volatility and information on the companies included in the tracking portfolio

5.1 Avarage spread per option contract

This table overviews the expected cost of buying and selling options contracts during the specified time interval. The average bid-ask spread per option contract, as depicted over the simulated time period, is shown in Table 5.1. A complete list of average bid-ask prices over the simulated period can be seen in Appendix B.

Table 5.1: Average spread for the 10 most accruing companies

Name	Average
ABB	10%
Alfa Laval	11%
Assa Abloy	15%
Atlas Copco A	14%
Boliden	10%
Electrolux	10%
Ericsson	15%
Evolution	13%
Getinge	8%
Handelsbanken	18%

5.2 Normality in returns

The returns of each constituent stock are distributed in a way that is similar to a normal distribution. Figure 5.1 represents the histogram over the returns from each component in the OMSX30 index. The histogram of the index returns is displayed in Figure 5.2, which also shows similarities to a normal distribution.

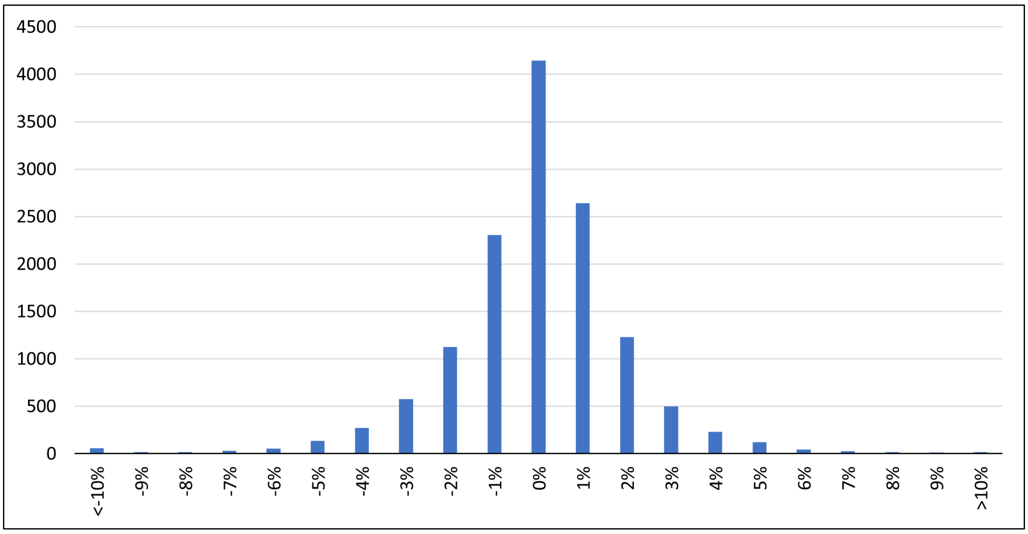

Figure 5.1: Histogram over returns from each OMXS30 component

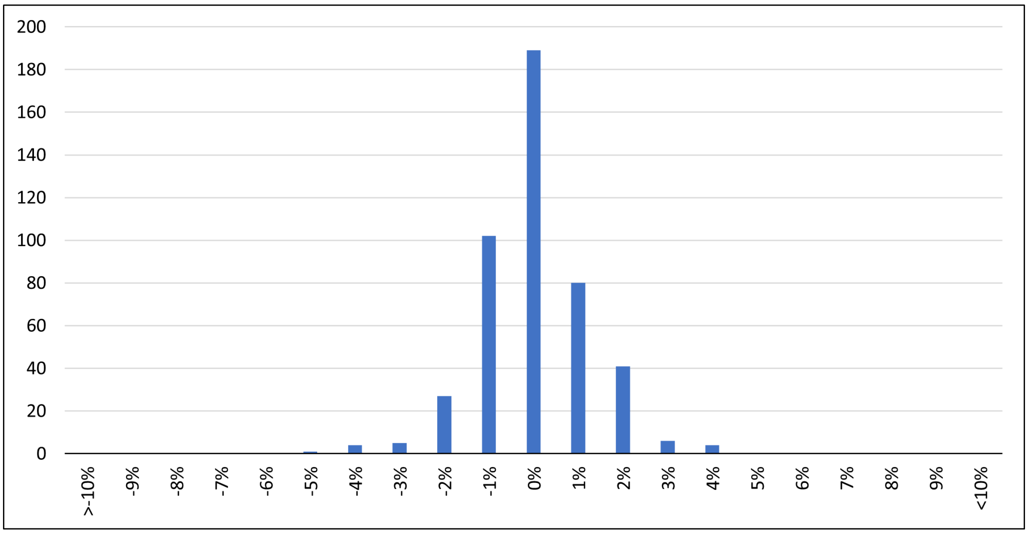

Figure 5.2: Histogram over returns from OMXS30 index

5.3 Implied volatility

Figure 5.3 presents a graphical representation of the daily ATM implied volatility with six months to maturity observed during the testing interval from August 30th, 2021, to March 24th, 2023. The red and blue lines represent the Markowitz implied volatility (MIV) and index option implied volatility (IOIV), respectively.

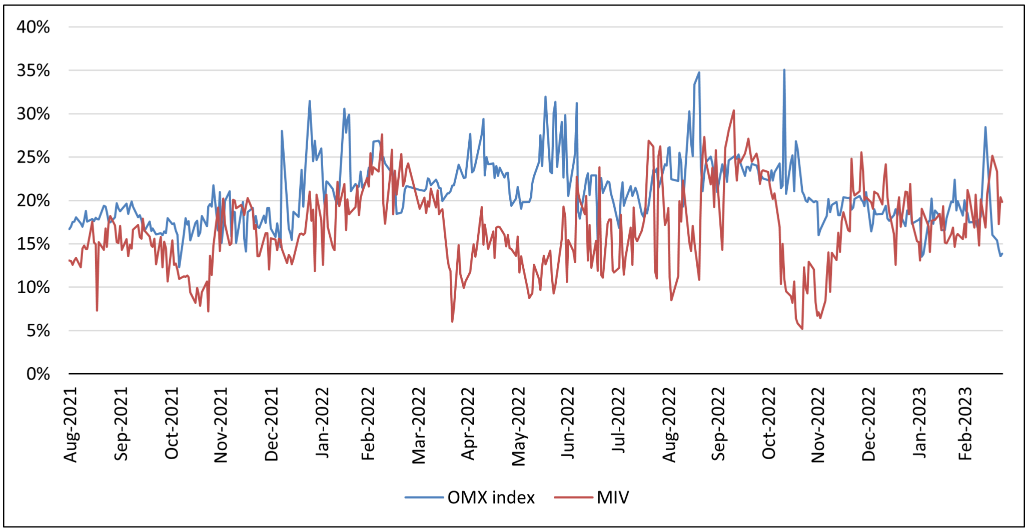

Figure 5.3: Index option implied volatility (IOIV) versus Markowitz implied volatility (MIV)

5.4 Tracking portfolio

The frequency of each OMXS30 component's participation in the tracking portfolio is shown in Figure 5.4. The graph summarises how frequently each constituent stock is included in the portfolio.

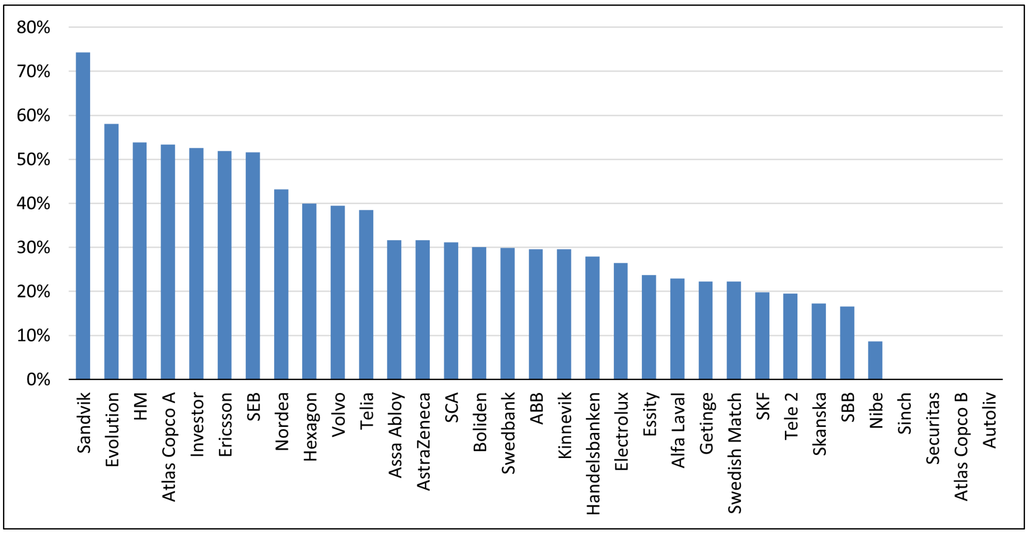

Figure 5.4: Number of time each OMXS30 component was included in the replicating portfolio

Figure 5.5 represent the R^2 values and R^2_{Adj}, from the replicating portfolio. The R^2 and R^2_{Adj} numbers indicate how well the replicating portfolio model fits the data and takes values between 0 and 1.

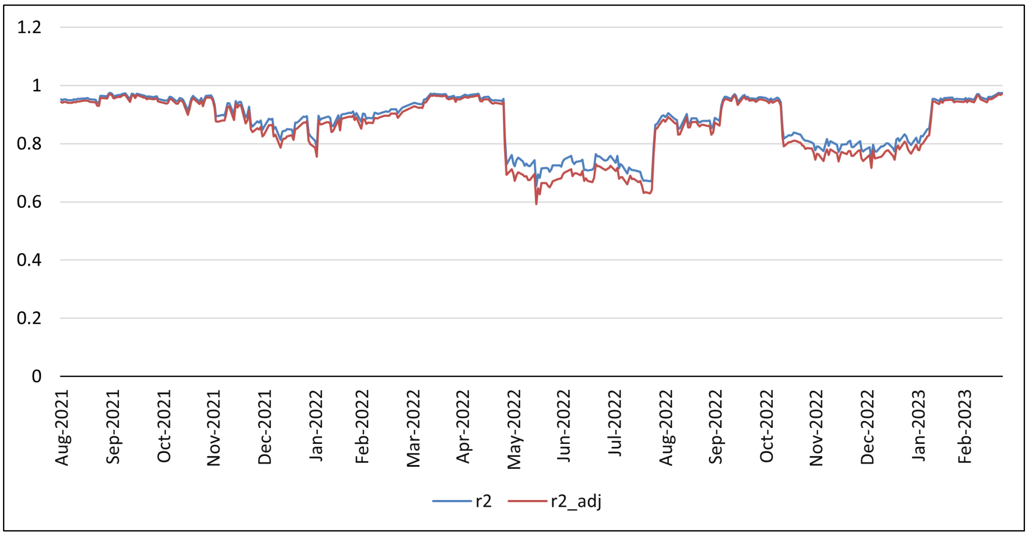

Figure 5.5: R^2 and R^2_{Adj} from the multiple linear regression

Figure 5.6 represents the Root Mean Squared Error, RMSE, obtained from fitting the replicating portfolio. The value of RMSE represents how well the tracking portfolio can replicate the returns of the OMXS30 index.

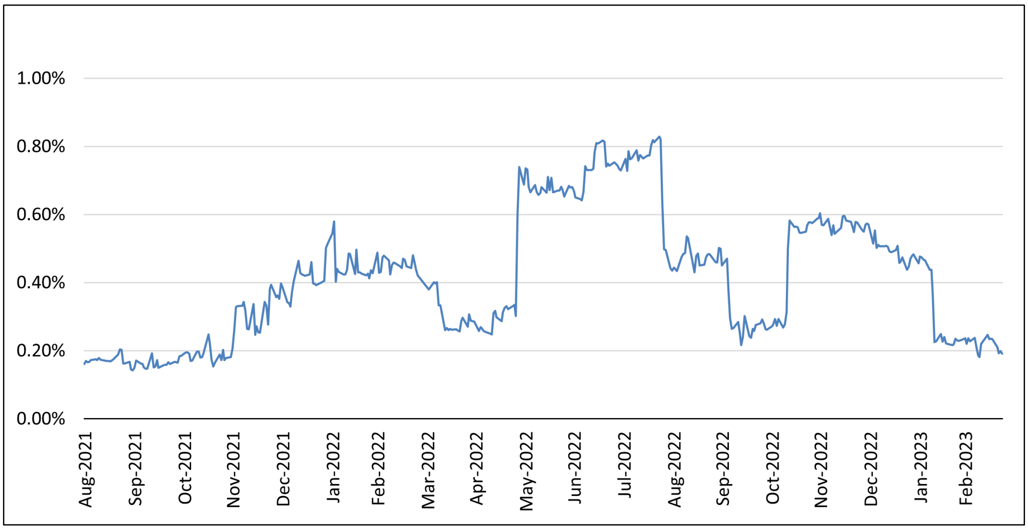

Figure 5.6: RMSE in the tracking portfolio

Figure 5.7 represents the annualized tracking error based on the lagging 90 days of returns. Tracking error represents how much the tracking portfolio is expected to deviate from the OMXS30 index on an annualized basis.

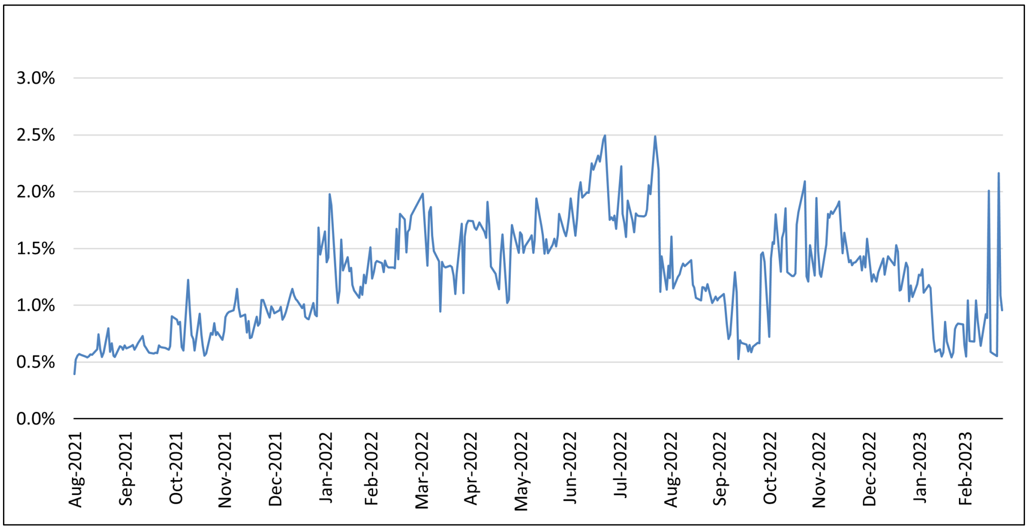

Figure 5.7: Tracking error in the tracking portfolio

5.5 Profitability

The mid price returns associated with holding the strategy until expiry is presented in Figure 5.8. The graph depicts two lines: the blue line illustrates the returns for the mid price when the portfolio is weighted with zero penalty, and the orange line displays the returns for the mid price when the portfolio is weighted with a constant value of 0.4. Figure 5.9 shows the returns for the full bid-ask spread for the strategy.

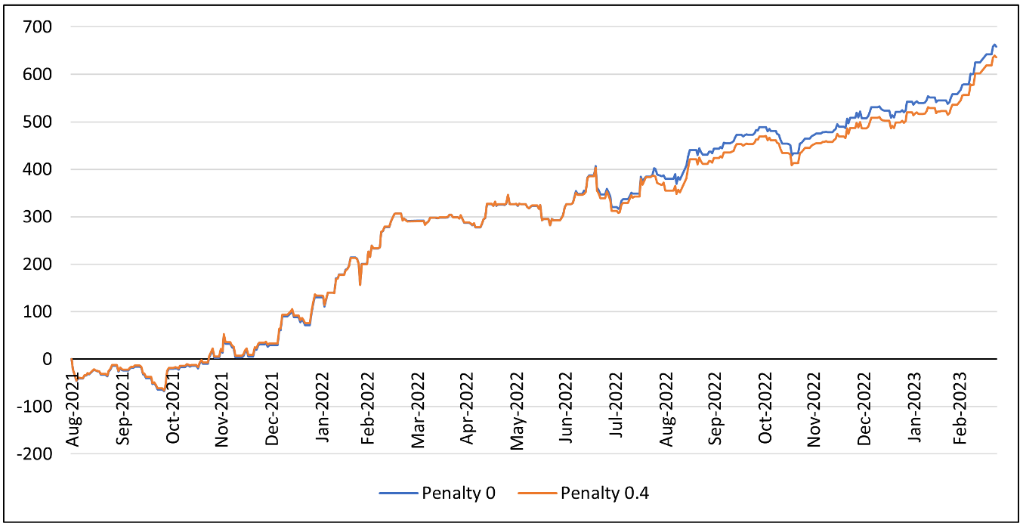

Figure 5.8: Returns for the mid price considering vega neutrality.

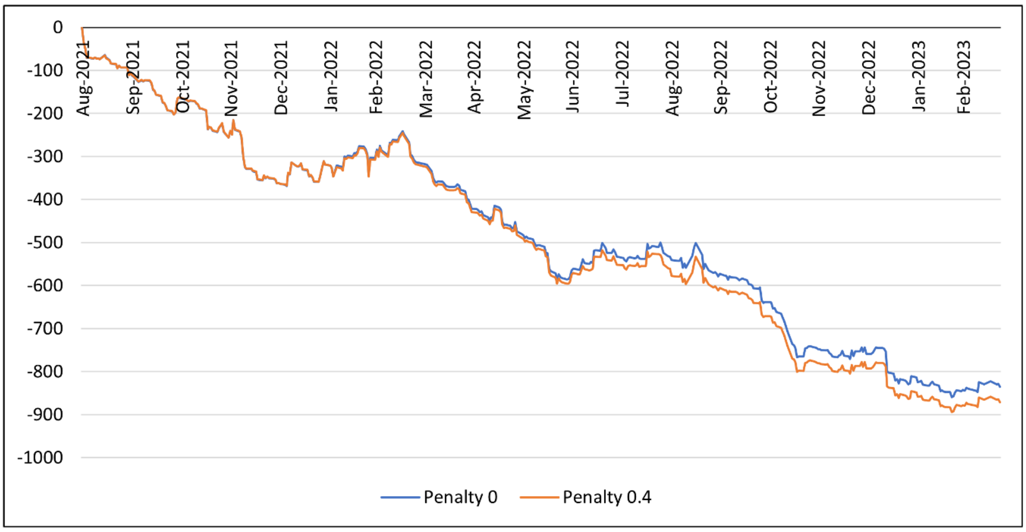

Figure 5.9: Return for full bid-ask spread considering vega neutrality.

Figure 5.10 illustrates the returns of holding equally weighted index straddles and replicating portfolio straddles in terms of mid-price. The various lines in the graph represent portfolios weighted with a penalty factor of 0 and 0.4, respectively. Figure 5.11 shows the returns for the full bid-ask spread for equal weights of index straddles and straddles on the replicating portfolio.

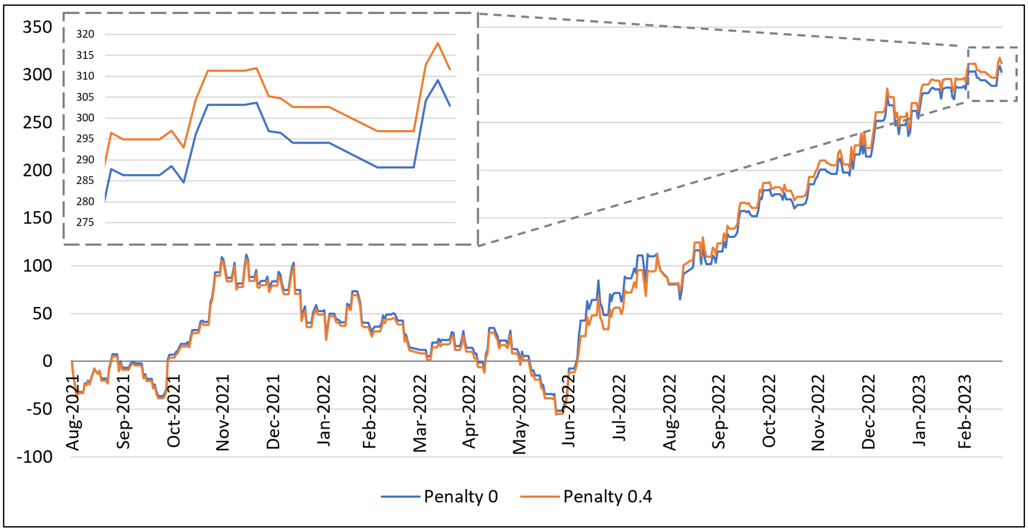

Figure 5.10: Return for mid price not considering vega neutrality.

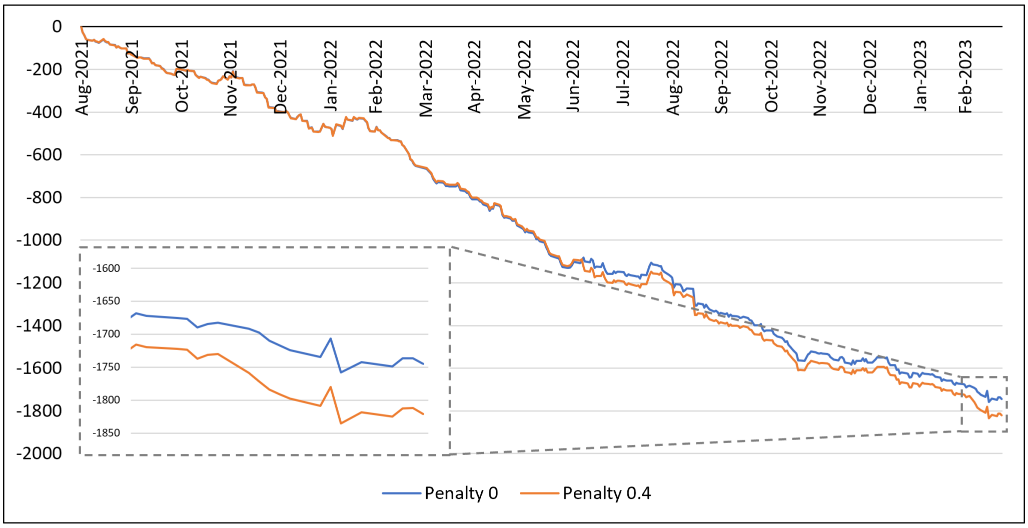

Figure 5.11: Returns for full bid-ask spread not considering vega neutrality.

5.6 Risk factors

The portfolio's net theta and vega, are presented in Figures 5.12 and 5.13, respectively. These figures provide an overview of the portfolio's overall exposure to volatility and time decay for the two strategies. The blue line represents the Greeks when the vega is hedged, and the red line represents the Greeks when vega is not hedged.

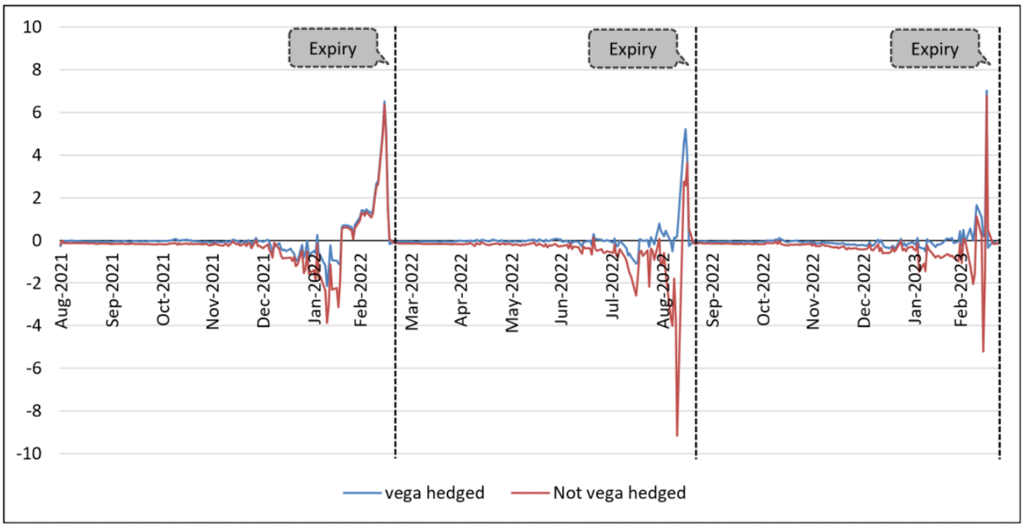

Figure 5.12: Portfolio net theta

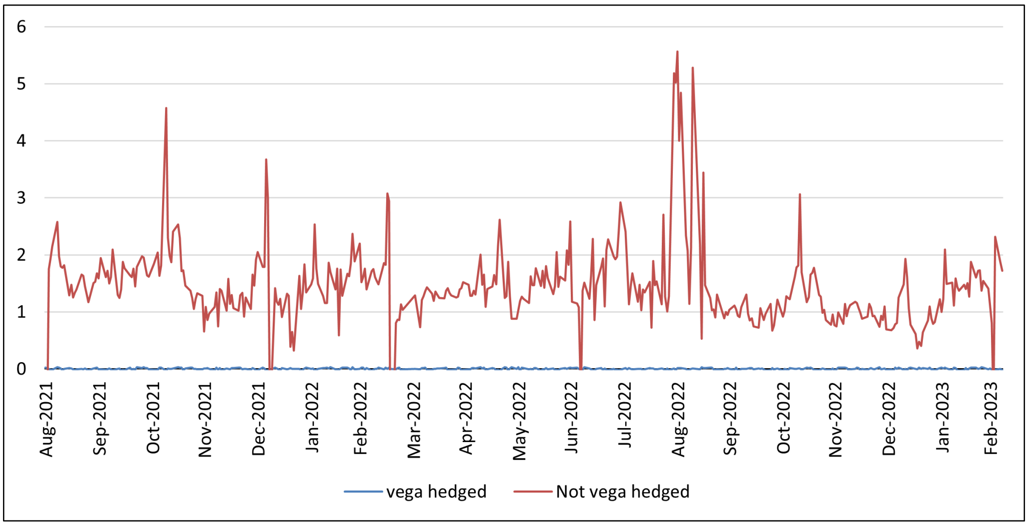

Figure 5.13: Portfolio net vega

6 Discussion

In this section the results of the dispersion strategy will be discussed. The results from the back-test using straddles will be discussed including the performance of the tracking portfolio, the profitability of executing the strategy, a reflection of the greeks and implied volatility. Furthermore, a comparison to the theory and a discussion of the implications for liquidity providers will be included.

6.1 Results from backtesting

6.1.1 Implied volatility

In this simulation, a similar trend in the index option implied volatility (IOIV), and Markowitz implied volatility (MIV) as observed by Magnusson (2013) and Jokela (2022) can be seen in Figure 5.3, where IOIV tend to trade at relatively richer volatility. However, there are cases where the trend breaks, and the MIV of the tracking portfolio trades at higher levels. In a non-naive strategy, one would either exit dispersion positions or change direction by selling volatility on the tracking portfolio and buying volatility on the index. Deng (2008) has demonstrated the profitability of naive strategies, which can be attributed to the limited and typically short-lived instances where IOIV trades above MIV on a tracking portfolio. Jokela (2022) also presents similar results in his research, where he investigates the trend in IOIV and MIV for the OMXS30 index.

6.1.2 Normality in returns

The OMXS30 index's returns, together with the components, follow a distribution that resembles a normal distribution. Thus the assumption of normally distributed returns in the Black and Scholes model, discussed in the theoretical background, is fulfilled.

6.1.3 Tracking portfolio

The companies picked by the forward selection to best explain the variation in OMXS30 can be seen in Figure 5.4. The percentage on the Y-axis represents the number of times each of the companies was picked to be in the replicating portfolio. It is not surprising to see that some of the largest components in the OMXS30 index are picked more regularly. The selection process, which is primarily driven by correlation rather than weight in the index, does not always guarantee this outcome. Being a larger part of the index, however, automatically means that a larger part of the returns will be reflected in the index, which also seems to be reflected in the correlation factor of these companies.

As shown in Figure 5.5, the tracking portfolio generates consistently high R^2_{Adj} and R^2 almost throughout the simulation. Furthermore, a 90-day lagging tracking error ranging from a high of almost 2.5% to a low of around 0.5% on an annualised basis can be observed in Figure 5.7. This figure represents what the expected deviation from the index being tracked would be on an annual basis, given that the correlation recorded during this 90-day period stays the same. As index funds generally have a tracking error between 1 and 2%, the values of the tracking error for the tracking portfolio are considered to be within an acceptable range [18].

Figure 5.6 shows the root mean squared error for the tracking portfolio in each timestep. As can be seen, the RMSE indicates an error between 30 and 80 basis points. This is in line with what other studies, such as Magnusson (2013), have received. However, a 30-80 basis point deviation could be a cause of concern worth keeping in mind.

6.1.4 Profit and loss

When trading options, the highest cost is generally the bid-ask spread. Hence, a smaller bid-ask spread should give a substantial benefit to the net returns over time as less money is being paid for the transaction [23]. During the investigated time period, the returns generated by running the backtest with full bid-ask spread (as shown in Figure 5.9) are found to be less profitable compared to the returns generated with the mid-price (as shown in Figure 5.8). This is to be expected as everything, but the added transaction cost, is held constant. These results are also confirmed in the study by Jokela (2022). What was unexpected, however, was how unprofitable it was to actively trade in the options contracts according to the proposed strategy. The stable and consistently high returns that can be observed in Figure 5.8 are replaced by an equally consistent loss when including the bid-ask spread in the simulation. The inclusion of a penalty factor for companies with larger bid-ask spreads does not seem to provide any substantial improvements to the overall performance of the model either. Very similar results can be observed in the strategy where 100 SEK is shorted on OMXS30 index and 100 SEK is bought on the replicating portfolio Figure 5.10 Figure 5.11. The cause of the negative returns does seem to be related to the frequency in which transactions in the options contracts are made. A less frequently re-weighted tracking portfolio would greatly decrease the number of transactions being made in the options contracts while still maintaining an adequate level of correlation, as was proven by Magnusson (2013). In contrast to the aforementioned strategy of buying and selling straddles to maintain vega neutrality, Figure 5.10 and Figure 5.11 illustrate the returns when holding equally cash-weighted index straddles and replicating portfolio straddles for the mid-price and full bid-ask spread respectively. The results demonstrate that, at this particular time frame, not hedging volatility risk is less profitable for both the mid-price and the full bid-ask spread. This confirms the theory that dispersion trading is a useful tool to protect traders against risks in implied volatility [13]. The average bid-ask spread for each option contract during a simulated time period is shown in detail in Appendix B. In contrast, an overview of the ten most occurring companies in the tracking portfolio is shown in Table 5.1.

6.1.5 Risk factors

As the delta is hedged, the graph is excluded from the result since it is zero throughout the simulation.

As theta and vega typically influence each other through volatility, one can see fluctuations in Figure 5.12 between the two lines. When volatility increases, the option price increases (affecting vega) while the time value decreases (affecting theta). Conversely, when volatility decreases, the option price decreases (affecting vega) while the time value increases (affecting theta) [17]. Theta stays close to zero almost throughout the whole simulation, but for the six-month options, one can see that it decay and peeks when it's close to expiration. The closer an option is to expiration, the faster the option price tends to decrease due to time decay. This means that options with a shorter time to expiration have a larger theta than those with a longer time to expiration.

When vega is near zero, the option's value is not very sensitive to changes in implied volatility [17]. In the strategy where vega is hedged, the simulation results in Figure 5.13 demonstrate effective vega hedging throughout the entire duration. This is evidenced by the consistent and controlled vega exposure depicted in the graph. In other words, changes in implied volatility do not significantly impact the strategy for this specific simulation as the value of vega is being hedged and subsequently close to zero. When vega is not hedged, on the other hand, a lot of underlying movements are seen in the implied volatility. If implied volatility increases, it suggests that market participants anticipate larger price swings in the underlying asset. This can result in higher option prices due to the increased uncertainty and perceived risk [17]. Conversely, when implied volatility decreases, it indicates a reduction in expected price fluctuations, leading to lower option prices.

6.2 Implications for liquidity providers

Using the proposed dispersion strategy offers an effective way of hedging vega risk consistently throughout a simulated period. As previously stated in section 2, it has been proven by other researchers such as Choi, C. Y. (2008) that vega can be hedged profitably in a dispersion trading strategy. Looking at the strategy's returns, however, it becomes unprofitable and expensive to execute when considering the full bid-ask spread in the options contracts. Although, the actual impact on profitability from paying the bid-ask spread may be better than it seems, as it is uncommon to pay the full bid-ask spread in reality. However, given the strong negative returns of paying the total bid-ask price and the minor improvements when penalising companies with larger than average bid-ask spreads, it is safe to say that even a smaller bid-ask spread would be unprofitable. Furthermore, it is worth considering carrying out the strategy on a large scale probably will have an impact on the prices in the market, as the market participants will notice that there is substantial demand on the bid/offer side on the included options contracts. This will likely have the effect of ”pushing the prices away”, thus adding a hidden transaction cost to the strategy. On top of this, a commission will also be paid on each transaction.

7 Conclusion

In this chapter, an evaluation of the results will be concluded. Positive effects and drawbacks of the strategy will be discussed as well as future work.

The simulation proved that the vega risk could be well hedged using the proposed strategy. Under the condition that no transaction costs are paid during the simulation, the hedging method would also result in positive results. Although the returns from the strategy are notably high, taking into account the substantial transaction costs renders the strategy unprofitable during the simulated period. The large transaction costs are the result of the weights in the tracking portfolio being re-calculated on a daily basis. A less frequent re-balancing of the weights in the tracking portfolio would result in lower transaction costs but would likely result in a worse hedge and lower correlation to the index. The profitability when trading at mid spread is a consequence of expensive index volatility being hedged with cheap single-stock volatility.

7.1 Future Work

Further studies on how vega can be hedged in general, and in the Swedish market in particular, would help to make options more liquid and thus a viable tool for market participants. An interesting topic might be how to incorporate options from other, more liquid markets and how that can be used to hedge risks associated with trading in Swedish options. Furthermore, investigating the frequency in which the weights of the tracking portfolio should be re-calculated would be interesting to investigate to decrease the transaction costs.

7.2 Final Words

The outlined approach can be utilized by SEB to hedge vega risk while trading options. However, it should be noted that this strategy has not shown to be adequate for generating profits when considering transaction costs over the selected time period.

References

[1] Alexander, C. *Market risk analysis. Volume 2*. Practical financial econometrics (1st edition), 2008.

[2] Bakshi, G., Kapadia, N., and Madan, D. "Stock return characteristics, skew laws and the differential pricing of individual equity options". In: *The Review of Financial Studies* (2003), 16 (1): 101–143.

[3] Beasley, J. E., Meade, N., and Chang, T. J. "An Evolutionary Heuristic for the Index Tracking Problem". In: *European Journal of Operational Research* (2002), 2003(148):621–643.

[4] Berk, J. and Demarzo, P. *Corporate Finance*. Pearson Education Inc, Boston, second edition edition, 2011.

[5] Björk, T. *Arbitrage Theory in Continuous Time, third edition*. Oxford University Press, 2009.

[6] Blume, M. E. and Edelen, R. M. *On Replicating the SP 500 Index*. The Wharton School University of Pennsylvania, 2002.

[7] Choi, C. Y. "Analysis and Development Of Correlation Arbitrage Strategies on Equities". In: *Working paper* (2008).

[8] Choudhry, M. *The Bond Money Markets, Strategy, Trading, Analysis*. Securities Institution Professional Reference Series, 2001, pp. 948–965.

[9] Deng, Q. "Volatility Dispersion Trading". In: (Jan. 2008).

[10] Driessen, J., Maenhout, P., and Vilkov, G. "Option-Implied Correlations and the Price of Correlation Risk". In: *Working Paper* (2005).

[11] Fama, E. F. "Efficient Capital Markets: A Review of Theory and Empirical Work". In: *The Journal of Finance* (1970), 25 (2): 383–417.

[12] Fatemi, B. and Krol, S. "Equity Dispersion". In: *Working paper* ((n.d)).

[13] FDAX Hunter. "Dispersion: A Guide for the Clueless". In: *Equity Derivatives Capital Structure Demolition* (2004). URL: `https://pdfcoffee.com/qdownload/dispersion-a-guide-for-the-clueless-pdf-free.html`.

[14] Ganatra, S. "Implementation of varaince swaps in dispersion trading strategies". In: *Master thesis, Cass Business School, London* (2004).

[15] Glosten, L. R. and Milgrom, P. *Bid, ask, and transaction prices in a specialist market with heterogeneously informed agents*. Journal of Financial Economics, 1985, 14(1985):71–100.

[16] Huang, X., Huang, Z., Ni, S. "Dispersion Trading and Its Implications for Volatility Estimation". In: *Journal of Financial Econometrics* (2017), 15(2), 175–202.

[17] Hull, J. C. *Options, futures, and other derivatives*. Pearson Education, 2018.

[18] Informa Statfacts. "Tracking Error". In: (2016). URL: `https://financialintelligence.informa.com/~/media/Informa-Shop-Window/Financial/StatFACTs/statfacts-trackingerror.pdf`.

[19] James, G., Witten, D., Hastie, T. and Tibshirani R. *Introduction to Statistical Learning - with Applications in R*. Springer Texts in Statistics (STS, volume 103), 2013.

[20] Jokela, A. "Dispersion Trading: Evidence from the Swedish Options Market". In: *Aalto University School of Business Finance* (2022).

[21] Krink, T., Mittnik, S., and Paterlini, S. "Differential evolution and combinatorial search for constrained index-tracking". In: *Ann Oper Res* (2009), 2009(172):153–176.

[22] Lisauskas, J. "Dispersion Trading in German Option Market". In: *Master Thesis, Tilburg University* (2011).

[23] Madhavan, A. *Market microstructure: A survey*. Journal of Financial Markets, 2000, 3(3):205–258.

[24] Magnusson, L. "Dispersion Trading: Construction and Evaluation". In: *Master thesis, KTH* (2013).

[25] Malkiel, B. G. "The Efficient Market Hypothesis and Its Critics. Journal of Economic Perspectives". In: *Journal of Economic Perspectives* (2003), 17 (1): 59–82.

[26] Maringer, D. and Oyewumi, O. “Index tracking with constrained portfolios”. In: *Intelligent Systems In Accounting Finance and Management* (2007), 2007(15):57–71.

[27] Markowitz, H. *Portfolio selection*. Journal of Finance, 1952, 7(1):77–91.

[28] Marshall, C. M. “Dispersion Trading: Emperical Evidence from U.S. Options Market”. In: *Global Finance Journal* (2008b), 2009(20):289–301.

[29] Marshall, C. M. “Isolating the systematic component of a single stocks (or portfolios) standard deviation”. In: *Technical report* (2008a).

[30] Martin, I. “Simple Variance Swaps”. In: *Working Paper* (2013).

[31] Maze, S. “Dispersion Trading in South Africa”. In: *Master Thesis, University of Cape Town* (2012).

[32] Morningstar. “Explainer: What is Tracking Error”. In: *Morningstar* (2021). URL: `https://shorturl.at/bpwFI`.

[33] Sant’Anna, L. R., Delgado de Oliveira, A., Filomena, T. P., Caldeira. J. F. “Solving the index tracking problem based on a convex reformulation for cointegration”. In: *Science Direct* (2018).

[34] Team, CFI. “Delta Hedging; A trading strategy that reduces the directional risk associated with the price movements of an underlying asset”. In: (2022). URL: `https://corporatefinanceinstitute.com/resources/derivatives/delta-hedging/`.

Appendix - Contents

A OMXS30 index simulated period

Table A.1: Companies in the OMXS30 index over the simulated period, and if they have options

Name	Options	Date included/excluded
ABB	Yes	-
Alfa Laval	Yes	-
Assa Abloy	Yes	-
Atlas Copco A	Yes	-
Atlas Copco B	No	-
Autoliv	No	-
AstraZeneca	Yes	-
Boliden	Yes	-
Electrolux	Yes	-
Ericsson	Yes	-
Essity	Yes	-
Evolution	Yes	-
Getinge	Yes	-
Handelsbanken	Yes	-
HM	Yes	-
Hexagon	Yes	-
Investor	Yes	-
Kinnevik	Yes	-
Nibe	Yes	Included 2 January 2023
Nordea	Yes	-
Sandvik	Yes	-
SBB	Yes	Included 2 July 2022
SCA	Yes	-
SEB	Yes	-
Securitas	Yes	Excluded 1 July 2021
Sinch	Yes	Included 2 July 2021
Skanska	Yes	Included 1 July 2022
SKF	Yes	-
Swedbank	Yes	-
Swedish match	Yes	Excluded 1 January 2023
Swedbank	Yes	-
Telia	Yes	-
Tele 2	Yes	-
Volvo	Yes	-

B Average Bid-ask spread

Table B.1: Average bid-ask spread per company, evaluated daily when included in the tracking portfolio.

Name	Bid ask spread
ABB	10%
Alfa Laval	11%
Assa Abloy	15%
Atlas Copco A	14%
Atlas Copco B	N/A
Autoliv	N/A
AztraZeneca	10%
Boliden	10%
Electrolux	10%
Ericsson	15%
Essity	13%
Evolution	13%
Getinge	8%
Handelsbanken	18%
Hexagon	19%
HM	13%
Investor	12%
Kinnevik	13%
Nibe	19%
Nordea	15%
Sandvik	11%
SBB	12%
SCA	11%
SEB	12%
Sinch	N/A
Skanska	9%
SKF	16%
Swedbank	15%
Swedish Match	20%
Tele 2	11%
Telia	28%
Volvo	5%

TRITA-SCI-GRU 2023:316