UNIVERSITY OF TURKU
Turku School of Economics

Determinants of the Implied Volatility Function

Empirical Evidence from the Euro Stoxx 50 Option Market

Bachelor's thesis
in Accounting and Finance

Author:
Matias Tuominen

Supervisor:
M.Sc. Ville Kukkonen

06/05/2026
Turku

Student's statement regarding the use of Artificial Intelligence (AI) for preparing and/or writing this thesis:

☐ **I have not used any AI-based tools.**

☒ **I have used AI-based tools**. Their use is documented in the Appendix. The AI tools were used in a way that complies with academic integrity guidelines.

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin Originality Check service.

Bachelor's thesis

Subject: Accounting and Finance
Author: Matias Tuominen
Title: Determinants of the Implied Volatility Function
Supervisor: M.Sc. Ville Kukkonen
Number of pages: 30
Date: 06/05/2026

This thesis investigates the determinants of the implied volatility function (IVF) within the Euro Stoxx 50 index option market. While the Black-Scholes model remains an industry standard for option pricing, empirical evidence consistently highlights its failure to account for the variance of implied volatility across different strike prices, a phenomenon known as the volatility smile or smirk. The objective of this study is to examine how market momentum, historical volatility, and trading volumes affect the daily variations in the IVF.

The empirical analysis was conducted using a two-stage framework with daily data spanning from January 2016 to December 2025, strictly limited to options with a constant 30-day maturity. In the first stage, the daily IVF was modeled using a quadratic polynomial to extract the specific coefficients for its level, slope, and curvature. In the second stage, the variations in these coefficients were analyzed against the explanatory variables using time-series OLS regressions with Newey-West robust standard errors.

The empirical findings demonstrate that downward market momentum and historical volatility are the most significant drivers of the IVF. Both variables exhibit a strong positive correlation with the overall level of the IVF, and a significant negative correlation with its slope and curvature. Economically, these results provide robust evidence for asymmetric hedging demands and the "crash-phobia" phenomenon. As market uncertainty or downward momentum increases, investors disproportionately bid up the prices of out-of-the-money (OTM) put options to insure their portfolios against sudden downturns, effectively steepening the downward slope of the smirk. The thesis confirms that macroeconomic variables systematically shape option pricing and market dynamics.

Key words: Implied volatility function, Implied volatility surface, Black-Scholes model

Kanditutkielma

Oppiainen: Laskentatoimi ja rahoitus
Tekijä: Matias Tuominen
Otsikko: Implisiittisen Volatiliteettifunktion Selittävät Muuttujat
Ohjaaja: KTM Ville Kukkonen
Sivumäärä: 30
Date: 06/05/2026

Tämä tutkielma tarkastelee implisiittisen volatiliteettifunktion (IVF) selittäviä muuttujia Euro Stoxx 50 -indeksin optiomarkkinoilla. Vaikka Black-Scholes-malli on edelleen optioiden hinnoittelun yleinen standardi, empiirinen näyttö korostaa toistuvasti sen kykenemättömyyttä selittää implisiittisen volatiliteetin vaihtelua eri toteutushinnoilla, mikä on ilmiö, joka tunnetaan volatiliteettihymynä tai -vinoumana. Tämän tutkimuksen tavoitteena on tarkastella, kuinka markkinamomentum, historiallinen volatiliteetti ja kaupankäyntivolyymit vaikuttavat IVF:n päivittäisiin vaihteluihin.

Empiirinen analyysi toteutettiin kaksivaiheisella menetelmällä käyttäen päivittäistä dataa tammikuusta 2016 joulukuuhun 2025, rajaten tarkastelun tiukasti optioihin, joilla on 30:n päivän maturiteetti. Ensimmäisessä vaiheessa päivittäinen IVF mallinnettiin toisen asteen polynomilla sen tason, kaltevuuden ja kaarevuuden kertoimien eristämiseksi. Toisessa vaiheessa näiden kertoimien vaihteluita analysoitiin suhteessa selittäviin muuttujiin käyttämällä aikasarjojen OLS-regressioita ja Newey-Westin robusteja keskivirheitä.

Empiiriset tulokset osoittavat, että negatiivinen markkinamomentum ja historiallinen volatiliteetti ovat IVF:n merkittävimmät ajurit. Molemmat muuttujat korreloivat voimakkaan positiivisesti IVF:n yleisen tason kanssa sekä merkittävän negatiivisesti sen kaltevuuden ja kaarevuuden kanssa. Taloudellisesti nämä tulokset tarjoavat vankkaa näyttöä epäsymmetrisestä suojautumiskysynnästä ja niin sanotusta romahduspelko-ilmiöstä (crash-phobia). Markkinoiden epävarmuuden tai negatiivisen momentumin kasvaessa sijoittajat nostavat suhteettomasti out-of-the-money (OTM) -myyntioptioiden hintoja suojatakseen salkkujaan äkillisiltä kurssilaskuilta, mikä käytännössä jyrkentää vinouman alaspäin laskevaa kaltevuutta. Tutkielma vahvistaa, että makrotaloudelliset muuttujat muokkaavat järjestelmällisesti optioiden hinnoittelua ja markkinadynamiikkaa.

Avainsanat: Implisiittinen volatiliteettifunktio, Implisiittinen volatiliteettipinta, Black-Scholes-malli

CONTENTS

FIGURES

TABLES

1 Introduction

The Black-Scholes model (1973) laid a foundation for modern option pricing theory regarding European options. Since its introduction empirical evidence has found deficiencies in the model, most notable being the variance of implied volatility across different strike prices. This causes the so called "volatility smile" or "volatility smirk", which will be explained more thoroughly in later sections. Academic research has produced more sophisticated option pricing models, which are able to work around the flaws of the Black-Scholes model. Despite its deficiencies and existence of newer models the Black-Scholes model continues to act as an industry standard. Option prices are often quoted using implied volatilities provided by the Black-Scholes model (Daglish, Hull and Suo 2007). This gives relevance to the implied volatility function (IVF), which describes the level of implied volatility at different strike prices of a European option.

The flaws of the Black-Scholes model have been approached from multiple perspectives. Hull and White (1987) introduced the concept of stochastic volatility to option pricing. One of the inadequacies of the Black-Scholes model is the assumption of constant volatility in the underlying asset. Instead of a constant volatility, Hull and White assume that the volatility of the underlying asset also obeys a stochastic process independent of the asset price.

Stochastic volatility has been implemented to multiple option pricing models. Perhaps the most notable implementation of stochastic volatility is the Heston model. Heston (1993) provides a closed-form solution for a European option when the implied volatility and price of the underlying asset are correlated. Prior to this, models were forced to either use numerical methods or assume zero correlation. This allows the model to also account for the volatility skew.

Instead of diving deeper into the flaws of Black-Scholes model and focusing on alternative pricing models, this study aims to focus directly on the IVF and its determinants. This is a less researched topic compared to the previously mentioned ones, but still important given the frequent use of the Black-Scholes model and implied volatility surfaces. It can be characterized as a reduced-form approach that utilizes Black-Scholes due to its position as a simple industry standard. Focusing on the IVF allows for a simple approach to studying how economic variables affect the European option market.

1.1 Research question

This study aims to answer how market momentum, historical volatility, and market activity affect the form of the IVF of index options. By addressing this question, the study aims to approach the deficiencies of the Black-Scholes model in a less common way and contribute to how the IVF is interpreted. The research question is answered through empirical evidence from the Euro Stoxx 50 index option market. The empirical evidence is formed through first finding the form that best explains the IVF and then testing how the chosen determinants explain each coefficient of the form.

This sort of methodology was first introduced in acadmic literature by Peña, Rubio and Serna (1999), who examined the options of the Spanish IBEX-35. Since then similar empirical analysis has been conducted in other equity markets, which will be covered more closely in section 2.4. This study draws strong inspiration from Peña et al. and will act as a continuation to prior empirical evidence.

1.2 Structure

This thesis is organized in the following manner. Chapter 2 introduces the most relevant theory regarding the research question. This begins with an explanation of the Black-Scholes model and its assumptions that ultimately lead to the emergence of the volatility smirk. This allows the examination of the IVF and the theoretical explanations for the volatility smirk. This will include prior empirical evidence on this topic. Chapter 3 describes the data and methodology used to conduct empirical analysis. It explains what the data consists of and the statistical tests used to determine the significance of economic variables as determinants for the IVF. Chapter 4 covers the results produced by the empirical analysis. It describes how the chosen variables affect the IVF. Chapter 5 summarizes the findings, compares them to theory and prior research and suggests topics for future research accordingly.

2 Theory

This section provides a theoretical framework that is closely tied to the empirical analysis in chapter 3. The chapter first explains the Black-Scholes model along with its assumptions. The Black-Scholes model allows examination of the IVF which is combined with theoretical explanations for the volatility smirk and prior empirical evidence.

2.1 Black-Scholes model

Originally Black-Scholes model was introduced for pricing stock options, but since then it has been applied to options for other asset classes, such as index options. The idea behind Black-Scholes model relies on setting up a riskless portfolio consisting of a position in both the option and its underlying asset. To remain risk-free the portfolio must be rebalanced constantly. (Black and Scholes 1973.) In a market where arbitrages don't exist the return of the portfolio must be the risk free rate, r (Hull 2021, 352).

2.1.1 Assumptions in the Black-Scholes model

The Black-Scholes model assumes ideal market conditions. The assumptions used in the Black-Scholes model are the following (Black and Scholes 1973):

1. The short-term interest rate, r is known and remains constant through time.
2. The price of the underlying asset follows a random walk with a lognormal probability distribution, where the volatility remains constant through time.
3. The underlying asset doesn't pay dividends or other distributions.
4. The option is a European option.
5. There are no transaction costs for the option or the underlying asset.
6. The option and the underlying asset are divisible, meaning it is possible to buy, sell or borrow any fraction of them.
7. Short selling is permitted with full use of proceeds.

In practice, many of the assumptions are frequently violated. Section 2.3 covers how the assumptions of the Black-Scholes model lead to variance in implied volatility across strike prices.

2.1.2 The Black-Scholes formula

The price of a European call or put option can be derived using the Black-Scholes model, which provides the following equations:

$$c = S_0 N(d_1) - Ke^{-rT} N(d_2),$$

and

$$p = Ke^{-rT} N(-d_2) - S_0 N(-d_1),$$

where

$$\begin{aligned} d_1 &= \frac{\ln(S_0/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}, \\ d_2 &= \frac{\ln(S_0/K) + (r - \sigma^2/2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}, \end{aligned}$$

c is the price of a European call option, p is the price of a European put option, S_0 is the current spot price of the underlying asset, K is the strike price of the option, r is the continuously compounded risk-free interest rate, T is the time to maturity in years, and $N(\cdot)$ is the cumulative standard normal distribution function. (Black and Scholes 1973.)

2.2 Implied volatility function

The IVF relates an option's implied volatility to its strike price, K, and time to maturity, T. While the standard Black-Scholes model assumes that volatility is a constant parameter, empirical evidence shows that implied volatility varies systematically across different degrees of moneyness. This phenomenon is often referred to the volatility smile due to its parabolic shape, which was common before the 1987 crash. More recent evidence suggests the shape of the IVF is more of a sneer or a smirk than a smile. (Dumas, Fleming and Whaley 1998.)

In practice, the Black-Scholes model is often used to determine the implied volatility. The implied volatility can be derived from both put and call options using the Black-Scholes. The Black-Scholes equation cannot be rearranged to calculate implied volatility analytically, but it can be derived using numerical methods. (Hull 2021, 363.)

One example of a numerical method that can be used to calculate an implied volatility using an option price is the Newton-Raphson method, a common method used for solving nonlinear systems of equations. If the price of either a call or a put option, $w \in \{c, p\}$, in the Black-Scholes formula is expressed as a function $f(S_0, K, r, T, \sigma^*)$, the implied volatility σ^* can be straddled using the Newton-Raphson method:

$$\sigma_{n+1} = \sigma_n - \frac{f(\sigma_n) - w}{f'(\sigma_n)}$$

where σ_n is the nth estimate of σ^*, and f' is the first derivative of $f(\sigma)$ with respect to σ. (Manaster and Koehler 1982.)

The implied volatility can also be approximated analytically. Corrado and Miller (1996) presented an analytical approximation for implied volatility using the following formula:

$$\sigma = \frac{\sqrt{2\pi}}{S_0 + X}\left(c - \frac{S_0 - X}{2} + \sqrt{\left(c - \frac{S_0 - X}{2}\right)^2 - \frac{(S_0 - X)^2}{\pi}}\right)\frac{1}{\sqrt{T}},$$

where $X = Ke^{-rT}$. The accuracy of this approximation depends on the strike price and maturity of the option and is most accurate when using at-the-money (ATM) options with more than three months to maturity.

The time to maturity aspect won't be covered in this study, since the empirical analysis will only be limited to the IVF derived from options with 30 days to maturity. In turn the moneyness of an option receives more focus as it is the only variable used in the IVF in this study. The moneyness of an option is determined by the current market price of the underlying asset and the strike price for that specific option. To account for the time value of money, moneyness is often defined relative to the forward price, F, of the underlying asset, rather than current spot price, S_0:

$$m = \frac{K}{F},$$

where m is moneyness, K is the strike price and $F = S_0 e^{(r-q)T}$.

Put-call parity is a relationship that exists between European put and call options with the same underlying asset, strike price and maturity. Like the Black-Scholes model, put-call parity also relies on the assumption that arbitrages don't exist. The put call-parity can be demonstrated with two portfolios that create the same cash flows in all scenarios. With the help of put-call-parity the following conclusion can be made:

$$c_{BS} - c_{mkt} = p_{BS} - p_{mkt},$$

where the subscripts BS and mkt denote prices derived from the Black-Scholes model and the actual observed market prices. This means the pricing error by the Black-Scholes model is the same for both put and call options with the same maturity, strike price, and underlying asset is the same if arbitrages don't exist, which also means the implied volatility is the same. The put-call-parity doesn't require assumptions about the probability distribution of the underlying assets future price. (Hull 2021, 453-454.)

The put-call parity is especially important in the empirical analysis as it implies that European put and call options with identical parameters result in the same implied volatility. This allows the examination of only one IVF, since put and call options lead to approximately the same IVF. It is also important to recognize that put call parity doesn't hold perfectly even in arbitrage free markets due to factors such as transaction costs, bid-ask spreads, and differences between borrowing and lending rates.

2.3 Theoretical explanations for volatility smirk

The form of the IVF can be approached from two perspectives which are both extensively researched areas: the probability distribution of the underlying asset and market dynamics. In this study the probability distribution is divided into three more subtopics: the non-normal distribution of stock returns, stochastic volatility, and leverage effect. These perspectives together create a simple theoretical framework for the failure of the Black-Scholes model and why the IVF usually takes the form of either a smile or a smirk.

2.3.1 Non-normal distribution of stock returns

The Black-Scholes model assumes the normal distribution of returns however, empirical evidence suggests stock returns do not follow a normal distribution. According to early evidence of Fama (1965) the probability distribution of American stock returns tend to have fatter tails than the normal distribution, meaning extreme outcomes are more probable.

This non-normal distribution of stock returns has been further supported by more recent evidence by Cont. This evidence also suggests that in addition to the probability distribution of intraday S&P500 index returns having fat tails, the distribution is not symmetrical, and large downward movements are more common than equivalent upward movements. In addition to normal distribution, the returns of stock market have been proposed to follow stable distributions, the Student distribution, hyperbolic distributions, normal inverse Gaussian distributions, and exponentially truncated stable distributions. (Cont 2001.)

The fat tails of the probability distribution of stock returns shifts more probability mass to extreme outcomes. This causes the Black-Scholes model to overprice options close to 100 % moneyness and underprice options further from it. Extreme outcomes are more likely than the Black-Scholes model assumes making the probability that a deep out of money option will be exercised more likely.

2.3.2 Stochastic volatility

As mentioned in the section 2.1.1 the Black-Scholes model assumes a constant volatility of the underlying asset until maturity. When volatility is stochastic the Black-Scholes model tends to over price ATM options and underprice deep out-of-the-money (OTM) and in-the-money (ITM) options. (Hull and White 1987.)

While stochastic volatility has been a major topic in development of modern option pricing models, it may not be the most relevant factor when examining options with shorter maturities. Stochastic volatility is more relevant in options with long maturity, while sudden price jumps are more relevant when pricing options with shorter maturities (Bates 2000). Options with short maturities are more dictated by short term volatility, while long maturities leave time for unexpected changes in volatility.

2.3.3 Leverage effect

Another explanation for the downward slope of the IVF is leverage effect, which refers to the negative correlation between the returns of a stock and its volatility. When the value of equity decreases the proportion of leverage in firm's value is increased. This causes the volatility of equity to increase in the process. (Christie 1982.)

The concept of leverage effect is consistent with the empirically proven downward slope of the IVF, but the magnitude of it is limited if all else is expected to remain constant. Steep downward slopes need further explanations to fully explain the form of the IVF.

2.3.4 Crash-phobia

The market crash of October 1987 was a turning point for the Black-Scholes model. Leading up to the crash the Black-Scholes model was more accurate in pricing options. Post-crash the Black-Scholes model started to underprice OTM puts more strongly. This was possibly due to a change in market dynamics, where market participants started to value OTM put options more in the fear of similar market crashes. (Rubinstein 1994.)

The Black-Scholes model fails to capture the market dynamics, since it doesn't take into account the risk preferences of investors. The Black-Scholes model prices options only based on expected values, while markets price options based on supply and demand.

Examinations of the performance of option pricing models utilizing stochastic volatility reveal they offer only a partial explanation to the volatility smirk (Bakshi, Cao and Chen 1997). The downward slope of the IVF can be partially explained through net buying pressure of OTM put options. OTM put stock index options face a lot of demand due to their ability to hedge away portfolio risk (Bollen and Whaley 2004). This fact is supported by the absence of similar demand for OTM put options for individual stocks (Bakshi, Kapadia and Madan 2003) (Bollen and Whaley 2004).

The higher price caused by the net buying pressure is due to the market makers' inability to hedge risk perfectly. Higher demand increases the option price by an amount proportional to the variance of the unhedgeable part. (Gârleanu, Pedersen and Poteshman 2009.)

2.4 Prior empirical evidence

The most notable prior empirical evidence for this study is provided by Peña, Rubio and Serna (1999), who studied options of the Spanish equity index IBEX-35. Similar empirical analysis has also been replicated by Beber (2001) on the Italian equity index Mib30 and Tanha and Dempsey (2015) on the Australian equity index ASX SPI 200. All studies focus on the determinants of the IVF using slightly varying methods to examine the determinants of the implied volatility function, but the results of the empirical analysis differ significantly.

Previous empirical analysis has utilized a form for the IVF that includes a level, a slope, and a curve. Peña (1999) fitted daily IVF's over a more than two year period and found the following quadratic function to be the most suitable:

$$\sigma = b_0 + b_1 X + b_2 X^2 + \epsilon,$$

where σ is the implied volatility and X is the degree of moneyness. The same form has also been utilized by Beber (2001) and Tanha and Dempsey (2015). This form contains three components: level (b_0), slope ($b_1 X$), and curvature ($b_2 X^2$).

Prior evidence suggests weak market momentum increases the overall level of the IVF and could slightly increase the curvature of the IVF (Peña et al. 1999). This suggests there could be demand for both hedging further downward movements in the market with OTM put options and speculating on a change of direction with OTM call options, increasing both ends of the IVF.

Similarly the historical volatility plays a determining role in the IVF. Intuitively historical volatility leads to a over all level on the IVF, but Peña et al. (1999) also found intraday volatility to be correlated with a lower degree of curvature. In turn Beber (2001) found that historical volatility causes asymmetry in the IVF.

The volume of options traded has been shown to correlate negatively with the slope of the IVF (Peña et al. 1999). This suggests higher trading volumes lead to an increased demand for hedging possible market crashes through OTM put options.

The bid-ask spread of index options has been shown to increase the curvature of the IVF significantly (Peña et al. 1999). A higher bid-ask spread could be accentuated at deep OTM or ITM moneyness levels increasing the price.

The time to maturity on index options has been the most dominant determinant of the IVF across all studies in different market, including Spain (Peña et al. 1999), Italy (Beber 2001), and Australia (Tanha and Dempsey 2015). Shorter maturities tend to lead to a higher curvature most likely caused by higher sensitivity to near-term jump risk (Peña et al. 1999). In turn, options with longer maturities are more dictated by long term development of the underlying asset and not sudden price jumps. The role of maturity in the IVF won't be covered in the empirical analysis of this study where time to maturity is strictly limited to approximately 30 days.

The so called Monday effect has been studied extensively. French (1980) introduced it by concluding that stock returns were on average lower during Mondays. Later on French and Roll (1986) further concluded that volatility does not scale linearly with calendar time, and return variance was much larger during trading hours. This topic is important since whether volatility from weekend carries over to Monday could affect the IVF. Prior empirical evidence suggests the form of the IVF is different on Mondays. On average the IVF has a lower curvature on Mondays. (Peña et al. 1999.)

Prior empirical evidence is conflicting around the determinants of the IVF proving further empirical research is relevant in this topic.

3 Data and methodology

The data set used for empirical analysis spans from January 1st 2016 to December 31st 2025. The data includes values only from trading days, and days with no values were removed from the data set. All data is collected from Bloomberg. It is important to note that the implied volatilities are derived from index options and are limited to the Euro Stoxx 50 equity price index, which may not be representative of other asset classes. Prior reserach suggests that the shape of the IVF significantly differs between stock index options and individual stock options (Bollen and Whaley 2004).

3.1 Implied volatility data

This study utilizes Euro Stoxx 50 price index implied volatilities provided by Bloomberg. The implied volatilities are from 9 different levels of moneyness: 80 %, 90 %, 95 %, 97.5 %, 100 %, 102.5 %, 105 %, 110 %, and 120 %. These levels capture the most liquid part of the IVF, but do not account for extreme deep OTM and deep ITM. The implied volatilities are limited to options with 30 days to maturity. The implied volatilities are calculated by Bloomberg using their LIVE equity engine available for stocks and equity indexes, which uses various methodologies for estimating the implied volatility at certain moneyness levels. The implied volatility time series consists of end-of-the day volatility levels, which are multiplied by one hundred for the empirical analysis.

3.2 Explanatory variables

The empirical analysis includes four explanatory variables: market momentum, historical volatility, trading volume of the index, and trading volume of Euro Stoxx 50 price index options. These variables are not expected to explain the IVF fully since the financial markets are rarely this simple.

Market momentum (MKT_t) is measured as the logarithmic ratio of the 60 trading day moving average of Euro Stoxx 50 price index relative to its current level.

$$MKT_t = \ln\left(\frac{\frac{1}{60}\sum_{s=t-60}^{t-1} P_s}{P_t}\right),$$

where P is the Euro Stoxx 50 price index, and t is current date. This variable captures whether the price index is above or below the 60 day moving average and the magnitude of that difference.

Historical volatility ($SIGMA$) is calculated using an annualized 20 trading day standard deviation of the Euro Stoxx 50 price index, using a one day lag. Similar to the implied volatility data the historical volatility is multiplied by one hundred.

$$SIGMA_t = \sqrt{252} \times \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(R_{t-i} - \bar{R})^2} \times 100,$$

where N is the size of the rolling window: $N = 20$, R is the daily logarithmic return: $R = \ln(\frac{P_t}{P_{t-1}})$, and P_t is the Euro Stoxx 50 price index.

The daily trading volume of the Euro Stoxx 50 index ($VMKT_{t-1}$) is used to capture the over all activity of the stock market. The trading volume is the value of all trades of the stocks in euros included in the Euro Stoxx 50 index on current day. Similar to historical volatility, this variable includes a one day lag. The variable is used as a proxy for general market liquidity and activity.

The volume of option contracts traded ($VOPT_{t-1}$) is used to capture the liquidity of the option market. Volume is measured as the logarithm of the number of options traded during the day. The logarithmic conversion is used to mitigate extreme outliers. This variable also includes a one day lag. The variable represents the trading activity and demand within the option market specifically.

3.3 Determining the right form for the IVF

The first stage of tests focuses on finding the correct structural form for the IVF. For this stage the IVF is fitted to the following models:

$$\begin{aligned}
&\text{Model 1}: \sigma = b_0 + \epsilon \\
&\text{Model 2}: \sigma = b_0 + b_1 X + \epsilon \\
&\text{Model 3}: \sigma = b_0 + b_1 U + b_2 D^2 + \epsilon \\
&\text{Model 4}: \sigma = b_0 + b_1 U + b_2 X^2 + \epsilon \\
&\text{Model 5}: \sigma = b_0 + b_1 U + b_2 X^2 + b_3 D + \epsilon \\
&\text{Model 6}: \sigma = b_0 + b_1 X + b_2 (X-1)^2 + \epsilon
\end{aligned}$$

where X is the degree of moneyness, U is a variable defined as $U_i = X_i$ if $X_i < 1$ and 0 otherwise, used to capture potential asymmetries on the left side of the volatility function, and D is a variable defined as $D_i = X_i$ if $X_i \geq 1$ and 0 otherwise, used to capture potential asymmetries on the right side of the volatility function. It is important to note that model 6 can also be rearranged to $\sigma = (b_0 + b_2) + (b_1 - 2b_2)X + b_2 X + \epsilon$. This is very similar to the form used by Peña et al. (1999), $\sigma = b_0 + b_1X + b_2X + \epsilon$. These two forms result in the same R^2 in the first two tests that will be covered next.

To account for the possibility of any logarithmic forms, the IVF is additionally fitted through the following models:

$$\begin{aligned}
\text{Model 7} &: \sigma = b_0 + b_1 \ln(X) + \epsilon \\
\text{Model 8} &: \sigma = b_0 + b_1 \ln(X) + b_2(\ln(X))^2 + \epsilon \\
\text{Model 9} &: \sigma = b_0 + b_1U + b_2(\ln(X))^2 + \epsilon \\
\text{Model 10} &: \sigma = b_0 + b_1U + b_2(\ln(X))^2 + b_3D + \epsilon
\end{aligned}$$

Two tests are completed on each of these models. The first test fits the model into a regression model for every day over the sample period. The models are ranked in this test according to average adjusted R^2 between each daily regression.

The second test combines all implied volatilities to a single OLS regression model. This enables the examination of the statistical significance of each model. For this test the null hypothesis (H_0) for each model is that all coefficient are equal to zero. In addition this test provides R^2 and AIC derived from the pooled regression, giving further insight to the applicability of each model.

The IVF form used in the second stage of tests is the model with the highest R^2 in the first stage assuming the p-value of the model is below a significance level of 0.05 in the pooled regression. Determining a standard form for the IVF enables the second stage of tests, which includes testing explanatory variables for coefficients included in the standard from.

3.4 Determinig how explanatory variables affect the shape of IVF

The second stage of tests focuses on the determinants of the IVF.

The effects of week day on the coefficients of the IVF are tested separately. This test utilizes coefficients (b_{it}) derived in the first stage and dummy-variables for each week day. The test determines whether any of the week day variables need to be incorporated in the latter tests. The following regression is ran for each coefficient (b_{it}) separately:

$$b_{it} = \beta_{\text{MO}}\text{MON}_t + \beta_{\text{TU}}\text{TUE}_t + \beta_{\text{WE}}\text{WED}_t + \beta_{\text{TH}}\text{THU}_t + \beta_{\text{FR}}\text{FRI}_t + \epsilon_t,$$

where MON_T, TUE_t, WED_t, THU_t, and FRI_t are dummy variables for each weekday.

Additionaly two Wald tests are employed for each coefficient, with two null hypotheses.

First null hypothesis states that all weekday coefficients are equal to each other: $\beta_{\text{MO}} = \beta_{\text{TU}} = \beta_{\text{WE}} = \beta_{\text{TH}} = \beta_{\text{FR}}$. The second null hypothesis focuses on the Monday effect specifically stating: $\beta_{\text{MO}} = (\beta_{\text{TU}} + \beta_{\text{WE}} + \beta_{\text{TH}} + \beta_{\text{FR}})/4$.

The determinants of the implied volatility function will be tested in a single regression model that is run for each coefficient (b_{it}) separately similar to the test regarding week days. The regression model is the following:

$$b_{it} = \beta_0 + \beta_1 MKT_t + \beta_2 SIGMA_{t-1} + \beta_3 VMKT_{t-1} + \beta_4 VOPT_{t-1} + \epsilon_t, \tag{1}$$

where MKT_t, $SIGMA_{t-1}$, $VMKT_{t-1}$, and $VOPT_{t-1}$ are as covered in section 3.2.

This model may include dummy variables for weekdays depending on the results of the test on weekdays.

The statistical null hypothesis (H_0) for each independent variable in the second stage of tests is that the coefficient is equal to zero, which implies that the specific variable has no effect on the coefficient in question. The standard errors for this test are reported using the Newey-West approach, introduced by Newey and West (1987), which is a heteroskedasticity and autocorrelation consistent (HAC) approach. The optimal number of lags chosen for the Newey-West estimator is chosen as 5, corresponding to a full trading week.

To avoid false conclusions of causality, the stationarity of each variable is ensured. This is performed with an Augmented Dickey-Fuller (ADF) unit-root test. The test is completed on both coefficients characterizing the IVF and explanatory variables. The ADF is a statistical test, in which the null hypothesis (H_0) assumes that the tested time series contains a unit root, which suggests it is not stationary (Brooks 2019, 288). In case an explanatory variable is non-stationary, a differentiation is needed to achieve stationarity.

The robustness of the results is tested by dividing the data into two equally long periods: period 1 and period 2. The explanatory variables for each coefficient (b_{it}) are tested separately for each period using the same OLS regression. The resulting coefficient (β_{jp}) from both periods are compared to determine whether the results are consistent across periods.

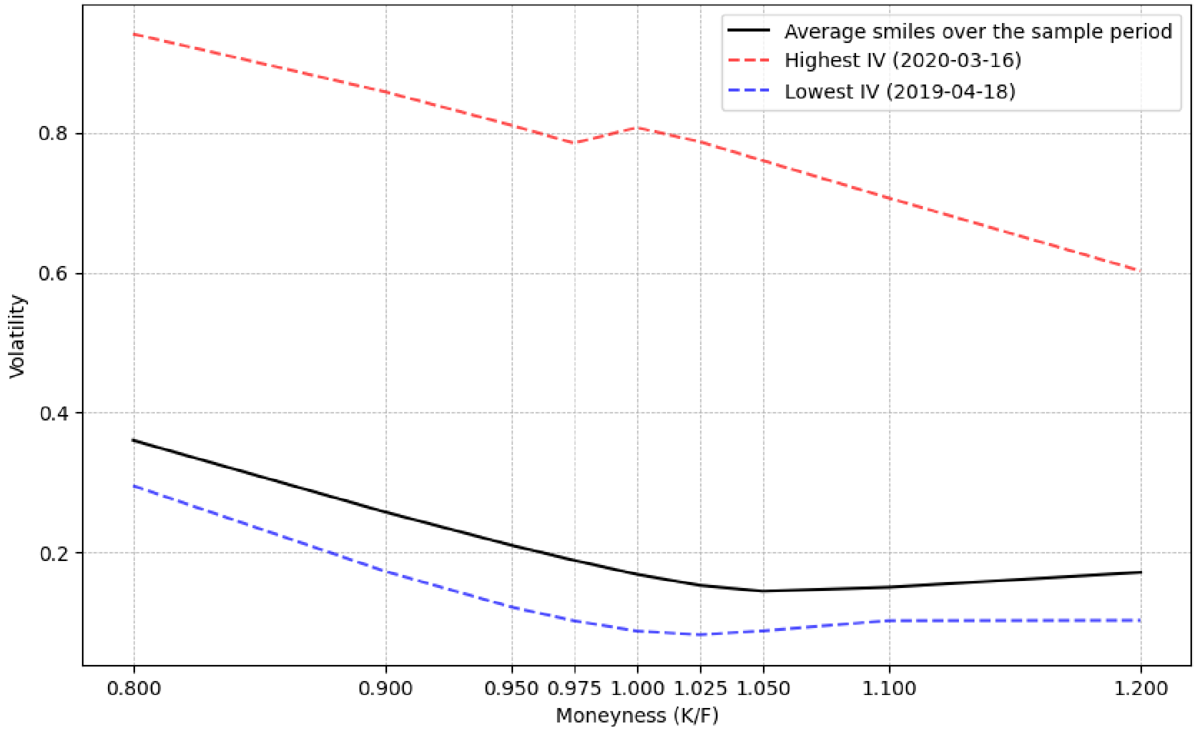

Figure 1: Average smile: STOXX50

4 Results

4.1 The correct form for IVF

Figure 1 presents visualizations of the IVF from the data set. It can be noticed that on average the IVF takes the form of a sneer, where there is a parabolic shape, but also a downward slope, as prior empirical evidence suggests. Additionally it can be noted that approximately the same form appears even in extreme situations, the lowest and highest volatility days in the ten year period, with the exception that during the highest volatility day the downward slope continues through the whole x-axis.

Table 1 presents the average adjusted R^2 for each model presented in section 3.3 starting from the highest value provided by the first test, the daily regression test. These values are derived from daily OLS regressions conducted over over the sample period, as introduced in section 3.3. The ten year data set consists of 2547 data points. The results reveal that model 6 has the highest explanatory power, with an average adjusted R^2 of 0.9757, suggesting model 6 best explains the form of the IVF.

Table 2 presents the results of the pooled regression test. The p-value for each model, except model 1, is below 0.001, making them statistically significant. These values support the daily regression test, since model 6 has again the highest adjusted R^2. Furthermore

Table 1: Function form: daily regression

Model	**Average adjusted R^2**
Model 6	0.9757
Model 8	0.9621
Model 9	0.8447
Model 10	0.8439
Model 7	0.6765
Model 2	0.6073
Model 4	0.4625
Model 5	0.4091
Model 3	0.2044
Model 1	0.0000

Note: Values represent the mean adjusted R^2 from daily cross-sectional regressions ($N = 2547$ days).

model 6 exhibits the lowest AIC of -63 921.14 providing further support for the model.

Table 2: Function form: pooled regression

Model	**Adj. R^2**	**AIC**	**F-statistic (p)**	**Params**
Model 6	0.5417	-63 921.14	< 0.001	3
Model 8	0.5357	-63 620.83	< 0.001	3
Model 10	0.4972	-61 794.78	< 0.001	4
Model 9	0.4866	-61 319.52	< 0.001	3
Model 7	0.4037	-57 887.40	< 0.001	2
Model 2	0.3689	-56 586.67	< 0.001	2
Model 5	0.3509	-55 940.59	< 0.001	4
Model 4	0.3325	-55 301.65	< 0.001	3
Model 3	0.2240	-51 848.18	< 0.001	3
Model 1	0.0000	-46 036.88	N/A	1

Note: Models are estimated using pooled OLS, where all data points are pooled into one OLS regression

Model 6 is chosen for the next stage of testing, since it has the highest R^2 in both tests, the daily regression and the pooled regression, and it is statistically significant. Model 6 is also able to explain different components of the IVF well, since it is simple a combination of a constant b_0, a slope b_1, and a curve b_2.

$$\text{Model 6} : \sigma = b_0 + b_1 X + b_2 (X - 1)^2 + \epsilon$$

This result is in line with previous research by Peña et al. (1999). Model 8 containing logarithmic features, resulted in also very similar results, with only slightly lower R^2 in

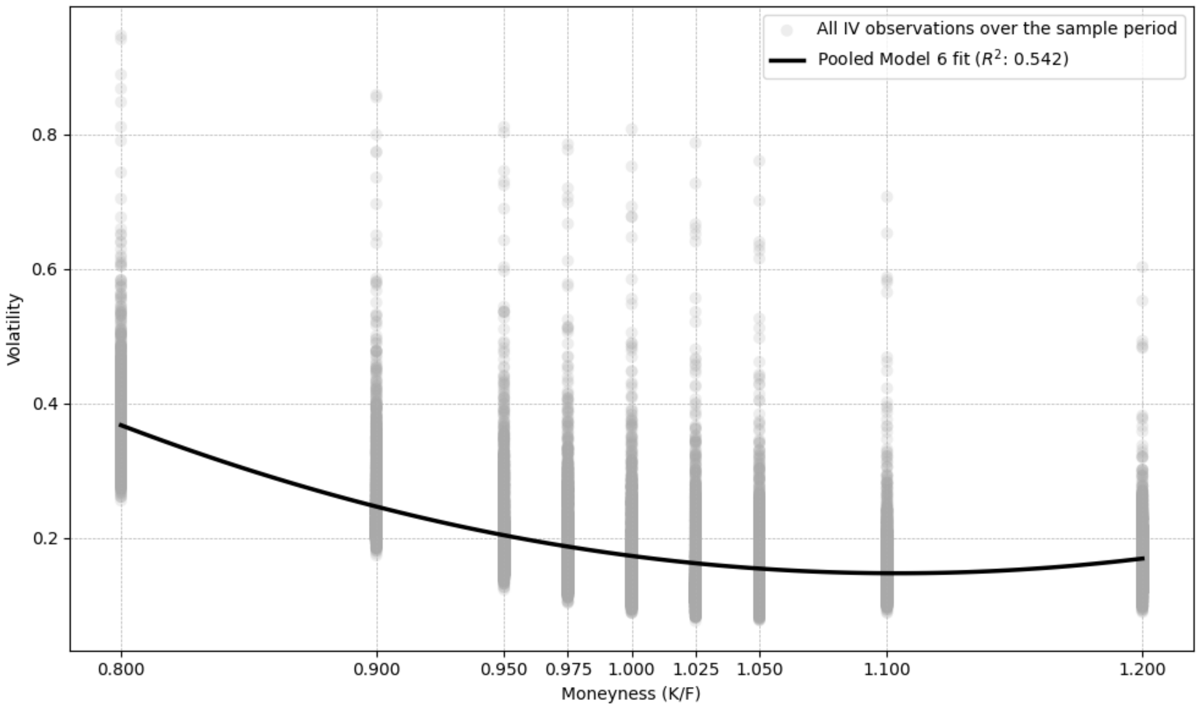

Figure 2: Pooled regression: STOXX50

both tests.

$$\text{Model 8} : \sigma = b_0 + b_1 \ln(X) + b_2(\ln(X))^2 + \epsilon$$

This is expected considering how similar model 6 and 8 are, but the tests suggest no logarithmic forms are needed.

Figure 2 visualizes the results of the pooled regression. Model 6 is presented with coefficients derived from the pooled regression. The individual data points show the consistency of the pattern.

4.2 Explanatory variables of the IVF

Table 3 presents the results of the seasonality test. Each coefficient is very similar during each week day suggesting week day doesn't significantly affect the IVF. Only exception is Friday which has a slightly higher degree of curvature. The p-value on the second Wald test ($p = 0.0830$) is not quite enough to make the observation statistically significant at the 5% level ($p < 0.5$). For other coefficients there doesn't seem to be any week day effects.

Since the day of the week doesn't significantly affect the form of the IVF, week day variables don't need to be included in the final test as explanatory variables. These

Table 3: Seasonality analysis of IVF parameters

Day of the week	b_0 **(Level)**	b_1 **(Slope)**	b_2 **(Curvature)**
Monday	0.6689***	-0.4952***	2.3769***
	(0.007)	(0.005)	(0.030)
Tuesday	0.6685***	-0.4954***	2.3679***
	(0.007)	(0.005)	(0.030)
Wednesday	0.6684***	-0.4956***	2.3627***
	(0.007)	(0.005)	(0.031)
Thursday	0.6691***	-0.4961***	2.3747***
	(0.007)	(0.005)	(0.031)
Friday	0.6675***	-0.4953***	2.4154***
	(0.007)	(0.005)	(0.032)
Wald Test (Monday vs Others)^a	0.8740	0.8806	0.8278
Wald Test (Joint Equality)^b	0.9955	0.9990	0.0830

Note: Standard errors in parentheses. Significance: *** $p < 0.01$.
^{a} p-value for the test $H_0 : \beta_{MON} = (\beta_{TUE} + \beta_{WED} + \beta_{THU} + \beta_{FRI})/4$.
^{b} p-value for the test $H_0 : \beta_{MON} = \beta_{TUE} = \beta_{WED} = \beta_{THU} = \beta_{FRI}$.

results differ from previous empirical evidence (Peña et al. 1999), where Monday has a lower degree of curvature compared to other week days.

Table 4: Augmented Dickey-Fuller test for stationarity

Variable	**ADF Statistic**	**p-value**
Level (b_0)	-6.7767***	0.0000
Slope (b_1)	-4.9414***	0.0000
Curvature (b_2)	-4.2795***	0.0005
MKT	-6.7960***	0.0000
$SIGMA$	-5.0652***	0.0000
$VMKT_{t-1}$	-6.1779***	0.0000
$VOPT_{t-1}$	-8.0457***	0.0000

Notes: The null hypothesis of the ADF test is that the series contains a unit root.
*** $p < 0.01$

Table 4 presents the results of the ADF test. All variables are significant at the 1 % level ($p < 0.01$). This allows the rejection of H_0 meaning the variables don't have a unit root, and they are stationary. This means no differentiation is needed to achieve stationarity. This also addresses the risk of spurious regression without differentiation.

Table 6 presents the results of the regression containing the determinants of the IVF. According to these results all tested variables effect the form of the IVF in some way at least at the 5% significance level. Out of four variables only the market momentum (MKT) and historical volatility ($SIGMA$) seem to have the most explanatory power

Table 5: Descriptive statistics of the explanatory variables

Variable	**Mean**	**Median**	**Min**	**Max**
MKT	-0.0060	-0.0100	-0.1700	0.4100
SIGMA	16.16	14.17	4.79	79.64
$VMKT_{t-1}$	15.7801	15.7718	8.9396	17.1152
$VOPT_{t-1}$	13.8035	13.7932	11.9658	15.3046

for the IVF, as their explanatory power for each coefficient is significant at the 1% level ($p < 0.01$).

All explanatory variables correlated positively with the level of the IVF (b_0). Together they reached an R^2 of 0.4723. The volume of options traded failed to reach statistical significance of 5% level ($p < 0.05$) while all other explanatory variables were significant at the 1% level ($p < 0.01$).

The effect of historical volatility is intuitive as past volatility can form future expectations, and the effect of market momentum is aligned with the theory of leverage effect, which according to volatility is supposed to decrease when the value of equity increases. This result suggests weak market momentum results in higher implied volatility or creates more demand for options across all strike prices. The effect of option trading volumes $VOPT_{t-1}$ is small in this sample, but it could mean high trading volumes in options are more driven by demand increasing the level of the IVF. The high trading volumes could indicate new information in the market that additionally causes volatility. The effect of stock market volume $VMKT_{t-1}$ is small but statistically significant.

Table 6: Determinants of the implied volatility function

	b_0 **(Level)**	b_1 **(Slope)**	b_2 **(Curvature)**
Intercept	-0.3871*	0.1551	3.8929***
	(0.218)	(0.175)	(0.907)
MKT	0.7939***	-0.4094***	-3.6916***
	(0.181)	(0.114)	(0.534)
SIGMA	0.6968***	-0.2904***	-1.9417***
	(0.086)	(0.060)	(0.350)
$VMKT_{t-1}$	0.0423***	-0.0360***	0.0552
	(0.014)	(0.011)	(0.063)
$VOPT_{t-1}$	0.0202*	-0.0027	-0.1517**
	(0.012)	(0.009)	(0.062)
Adjusted R^2	0.4723	0.2253	0.2831
Observations	1,493	1,493	1,493

Newey-West standard errors are reported in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

For slope (b_1) the effect of each explanatory variable is the opposite, meaning all explanatory variables have a negative correlation with the slope and are the causes of the downward slope. All variables except the volume of options traded are significant at the 1% level ($p < 0.01$). Together the variables have an R^2 of 0.2253 over b_1, slightly lower than for b_0.

The role of historical volatility is perhaps the most logical as past volatility could cause crash-phobia where OTM put options have a high demand for hedging purposes. The role of market momentum in determining the slope is less clear in prior empirical evidence, but the results suggest downward market momentum increases the downward slope. This could mean investors want to prevent further losses by hedging away downside creating demand for OTM put options.

For the slope stock market trading volumes seem to have a small, but statistically significant effect. The relationships between explanatory variables, level, and slope could be interpreted in an alternative way: higher implied volatility (level) leads to a steeped downward slope on the IVF, and therefore same variables that effect over all volatility, also correlate with slope.

All tested variables, except stock market volume that didn't reach statistical difference, the relationship to the degree of curvature is negative. All other variables were significant at least at the 5% level ($p < 0.05$), and market momentum and historical volatility were significant at the 1% level ($p < 0.01$).

It seems that same factors that cause a steeper downward slope, decrease the degree of curvature flattening the IVF. The results affecting degree of curvature don't have a clear explanation, suggesting the negative relationships are caused mostly by the higher downward slope.

Many of the observations in this study have been made in prior empirical analysis. For level similar effect of market momentum, historical volatility, and option trading volumes has been documented earlier by Peña et al. (1999). Research by Beber (2001) also supports the role of historical volatility. The effect of stock market volume on level hasn't been concluded in prior evidence. The effect of historical volatility on slope has been observed prior to this by both Peña et al. (1999) and Beber (2001). Additionally Tanha and Dempsey (2015) found market momentum to explain the downward slope partly.

Table 7 presents the results of the sub-period analysis. The dataset is divided around the start of the year 2021. The sub-period analysis strongly supports the impact of market momentum and historical volatility. There is some deviation between periods, but for each coefficient the sign of R^2 is the same and magnitude is roughly the same. With these

Table 7: Robustness Check: Sub-period Stability Analysis

	b_0 **(Level)**			b_1 **(Slope)**			b_2 **(Curvature)**		
	Full	**Per 1**	**Per 2**	**Full**	**Per 1**	**Per 2**	**Full**	**Per 1**	**Per 2**
Intercept	-0.3871	-0.0738	-0.9002	0.1551	-0.1216	0.5872	3.8929	2.7630	2.5625
MKT	0.7939	0.7962	0.9668	-0.4094	-0.4178	-0.5426	-3.6916	-2.5119	-4.4506
SIGMA	0.6968	0.7817	0.4525	-0.2904	-0.3437	-0.1379	-1.9417	-1.9372	-2.2214
$VMKT_{t-1}$	0.0423	0.0025	0.0795	-0.0360	0.0013	-0.0701	0.0552	-0.0139	0.1371
$VOPT_{t-1}$	0.0202	0.0408	0.0191	-0.0027	-0.0236	0.0020	-0.1517	-0.0101	-0.1273

Note: Period 1 covers the sample up to 2020-12-30. Period 2 begins on 2021-01-04.

results it is more likely that those variables have a consistent effect on the IVF.

The previous observations about stock market and option market volumes are heavily questioned by the sub-period analysis. In addition to the coefficients being close to zero, the coefficients switch signs between periods for some coefficients. This limits the conclusions we can make for trading volumes from the analysis, since the results do not seem to be consistent across periods.

5 Conclusions

This thesis studied the effects of four variables on the implied volatility function (IVF) of the Euro Stoxx 50 price index: market momentum, historical volatility, stock trading volume, and option trading volume. This was carried out utilizing a two-stage empirical framework. First the general form of the IVF was tested with two regressions: a daily OLS regression for each data point and a pooled OLS regression consisting of all data points. These tests together resulted in a quadratic form that captures the IVF's level, slope and curvature. In the second stage the explanatory variables were tested on the coefficients derived from the daily regression. This was done utilizing an OLS regression and Newey-West standard errors. The robustness of the results was further tested with a sub-period analysis, that provided information into whether results stayed consistent across periods.

The results of the empirical evidence shows there is a clear relationship between the IVF and the explanatory variables. Out of the four tested variables market momentum and historical volatility seemed to be the most relevant ones. The effects of stock market and option market volume were statistically significant in the initial test, but proved to be less clear in the sub-period analysis due to variance in results between periods. The empirical results suggest downward market momentum and historical volatility are positively correlated with the level of the IVF and negatively correlated with the slope and the curvature of the IVF. The empirical results also show that all variables that correlate positively with the level of the IVF tend to correlate negatively with the slope and the curvature of the IVF.

Economically, these findings provide strong support for the existence of a volatility sneer, most likely caused by asymmetric hedging demands within the stock index option market. This is supported by the strong explanatory power of historical volatility. Additionally the strong explanatory power of market momentum shows crash-phobia is not just driven by past volatility, but also downward market momentum. From a practical perspective, understanding these dynamics is highly valuable for all market participants, as recognizing how economic variables systematically affect the IVF allows practitioners to better anticipate market shifts and adjust their strategies. The findings illustrate that the volatility smirk is not a static error in the Black-Scholes model, but a dynamic phenomenon that can be approached from the perspective of economic factors affecting it.

Future research could expand the scope to include a more in-depth theoretical analysis between the variables and different components of the IVF. This study alone is inadequate for explaining whether the observed correlations are rooted in causality. The empirical analysis in this study could also be continued with deeper statistical methods such as

Granger causality tests, which have been utilized in prior empirical research. The relationships between level, slope and curvature could be examined more closely to give more context to the received results. The used data could be expanded to include different maturities, and more explanatory variables, such as bid-ask spreads of options or risk free rate.

References

Bakshi, G., Cao, C. – Chen, Z. (1997) Empirical performance of alternative option pricing models. *The Journal of Finance*, vol. 52 (5), 2003–2049.

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

Bates, D. S. (2000) Post-'87 crash fears in the s&p 500 futures option market. *Journal of Econometrics*, vol. 94 (1-2), 181–238.

Beber, A. (2001) Determinants of the implied volatility function on the italian stock market. Tech Report 10, ALEA.

Black, F. – Scholes, M. (1973) The pricing of options and corporate liabilities. *Journal of Political Economy*, vol. 81 (3), 637–654.

Bollen, N. P. B. – Whaley, R. E. (2004) Does net buying pressure affect the shape of implied volatility functions? *The Journal of Finance*, vol. 59 (2), 715–754.

Brooks, C. (2019) *Introductory Econometrics for Finance*. Cambridge University Press, 4th edn.

Christie, A. A. (1982) The stochastic behavior of common stock variances: Value, leverage and interest rates. *Journal of Financial Economics*, vol. 10 (4), 407–432.

Cont, R. (2001) Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, vol. 1 (2), 223–236.

Corrado, C. J. – Miller, T. W. (1996) A note on a simple, accurate formula to compute implied standard deviations. *Journal of Banking & Finance*, vol. 20 (3), 595–603.

Daglish, T., Hull, J. – Suo, W. (2007) Volatility surfaces: theory, rules of thumb, and empirical evidence. *Quantitative Finance*, vol. 7 (5), 507–524.

Dumas, B., Fleming, J. – Whaley, R. E. (1998) Implied volatility functions. *The Journal of Finance*, vol. 53 (6), 2059–2106.

Fama, E. F. (1965) The behavior of stock-market prices. *The Journal of Business*, vol. 38 (1), 34–105.

French, K. R. (1980) Stock returns and the weekend effect. *Journal of Financial Economics*, vol. 8 (1), 55–69.

French, K. R. – Roll, R. (1986) Stock return variances: The arrival of information and the reaction of traders. *Journal of Financial Economics*, vol. 17 (1), 5–26.

Gârleanu, N., Pedersen, L. H. – Poteshman, A. M. (2009) Demand-based option pricing. *The Review of Financial Studies*, vol. 22 (10), 4259–4299.

Heston, S. L. (1993) A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, vol. 6 (2), 327–343.

Hull, J. – White, A. (1987) The pricing of options on assets with stochastic volatilities. *The Journal of Finance*, vol. 42 (2), 281–300.

Hull, J. C. (2021) *Options, Futures, and Other Derivatives*. Pearson, Harlow, England, 11th ed. edn.

Manaster, S. – Koehler, G. (1982) The calculation of implied variances from the black-scholes model: A note. *The Journal of Finance*, vol. 37 (1), 227–230.

Newey, W. K. – West, K. D. (1987) A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, vol. 55 (3), 703–708.

Peña, I., Rubio, G. – Serna, G. (1999) Why do we smile? on the determinants of the implied volatility function. *Journal of Banking & Finance*, vol. 23 (8), 1151–1179.

Rubinstein, M. (1994) Implied binomial trees. *The Journal of Finance*, vol. 49 (3), 771–818.

Tanha, H. – Dempsey, M. (2015) Do aussie markets smile? implied volatility functions and determinants. *Applied Economics*, vol. 47 (30), 3143–3163.

Appendix 1: Disclosing the use of AI

Generative AI has been used in the making of this thesis. The usage of AI is described below. The usage of AI has been meticulous and followed the AI guidelines provided by University of Turku.

Used AI model: Google Gemini 3.1 Pro

Operating phase: 1. source searching, 2. programming, 3. LaTeX formatting, 4. proofreading

1. Source searching

Generative AI was used to find relevant academic sources. The relevance of each source was evaluated by the author. The decision of whether a source was used as part of the text was made by the author. AI was not used to read, summarize, or draw conclusions from any sources.

Example prompt: "Give me the most relevant academic sources for stochastic volatility."

2. Programming

Generative AI was used to debug Python code or generate appropriate Python syntax for certain statistical methods. AI was especially used for statistical models from statsmodels.api library and data visualization from matplotlib library. The quality of syntax generated by AI was verified by the author using prior knowledge of Python and documentation from previously mentioned libraries.

Example prompt 1: "Here is my code: {code}, here is the error message: {error message}"

Example prompt 2: "How can I execute the augmented Dickey-Fuller test in Python?"

3. LaTeX formatting

Generative AI was used to format formulas in LaTeX and move tables from Python to LaTeX. All text inserted into LaTeX was proofread by the author as well as the compiled result. The mathematical logic of each formula was verified by the author. Each table in LaTeX was compared to the original table in Python, and all text in tables was written by the author.

Example prompt 1: "Make this formula more readable: {formula}"

Example prompt 2: "Write this table in LaTeX: {Python table}"

4. Proofreading

Generative AI was used to proofread parts of the text for grammatical errors or false information. The feedback was then analyzed by the author.

Example prompt: "Are there any grammatical errors or false information in this text: {text}"