



Handelshøyskolen BI

GRA 19703 Master Thesis

Thesis Master of Science 100% - W

Predefinert informasjon

Startdato:	16-01-2022 09:00	Termin:	202210
Sluttdato:	01-07-2022 12:00	Vurderingsform:	Norsk 6-trinns skala (A-F)
Eksamensform:	T		
Flowkode:	202210 10936 IN00 W T		
Intern sensor:	(Anonymisert)		

Deltaker

Navn:

Informasjon fra deltaker

Tittel *:

Navn på veileder *:

Inneholder besvarelsen
konfidensielt
materiale?: Nei

Kan besvarelsen
offentliggjøres?: Ja

Gruppe

Gruppenavn:

Gruppenummer:

Andre medlemmer i
gruppen:

OBX-VIX: The predictability of OBX returns and asymmetric relationship between implied volatility and returns

Master Thesis

By

Læknes, Bjørn Magne

&

Melsbø-Nilsen, Emil

MSc in Business *Major in Finance*

ABSTRACT

In this paper, we have constructed an implied volatility index for the Norwegian equity market (OBX-VIX). We have tested its properties and have used it to look at the asymmetric relationship between implied volatility and equity returns. The sample period for our analysis spans an extensive period, from 2007 to 2020, to capture multiple financial downturns. We find evidence that OBX-VIX can predict movements in OBX, and OBX is a predictor of movements in OBX-VIX.

Supervisor:

Paul Ehling

Acknowledgments

We would like to thank our supervisor, Paul Ehling, for providing valuable insights and guidance during this project. We would also like to thank Anders Holen, Head of Derivatives at Euronext Oslo Børs, for his insights regarding the topic and for contributing additional data for our thesis.

Table of Contents

1. Introduction	1
2. Research question	2
2.1 Testable hypothesis	2
3. Literature review	3
3.1 Risk-reward relationship	3
3.2 Black-Scholes and volatility	3
3.3 Leverage effect	5
3.4 Asymmetric Volatility Phenomenon	5
3.5 Implied Volatility Index and stock return	5
3.6 Volatility clustering and mean reversion	6
3.7 Previous work on implied volatility index in Norway	6
3.8 Quantile regression approach	6
4. The research methodology and design	7
4.1 Research design	7
4.2 Methodology	7
4.2.1 VIX construction	7
4.2.2 Multiple Linear Regression	10
4.2.3 Quantile Regression.....	11
4.3 Data collection	13
4.3.1 OBX option	13
4.3.3 OBX Total Return index and Risk-free rate.....	14
5. Results	14
5.1 Descriptive analysis	14
5.2 Correlation	15
5.3 Multiple Linear Regression	17
5.3.1 Model and results	17
5.3.2 Linear regression assumptions and our data.....	22
5.4 Quantile Regression: OBX-VIX as the dependent variable	22
5.5 Quantile Regression: OBX return as the dependent variable	25
5.6 Discussion and findings of empirical evidence	27
6. Conclusion	29
Bibliography	30
Appendix	33

Table 1: Overview of data used to construct the OBX-VIX index.	13
Table 2: Descriptive statistics of OBX-VIX and VIX and their respective returns.	15
Table 3: Correlation between the different variables across different sample periods.	17
Table 4: Multiple Linear Regression, a summary of statistics from the six different models.	19
Table 5: Quantile regression, summarizing statistics from quantile 0.1-0.9.	24
Equation 1: Black & Scholes option pricing model	4
Equation 2: The general formula to calculate VIX	8
Equation 3: Interval between strike prices in VIX calculation	8
Equation 4: Time to maturity in VIX calculation.....	8
Equation 5: Future price on underlying asset in VIX calculation	9
Equation 6: Volatility for near- and next-term options in VIX calculation	9
Equation 7: 30-day weighted average value of the volatility in VIX calculation	10
Equation 8: Multiple Linear Regression model	11
Equation 9: Conditional quantile function	12
Equation 10: Minimizing problem equation	12
Equation 11: Quantile regression model	12
Equation 12: Multiple Linear Regression with the first difference of OBX-VIX as the dependent variable.....	18
Equation 13: Quantile function for each quantile	23
Equation 14: Quantile regression with returns of OBX-VIX as the dependent variable	23
Equation 15: Quantile regression with returns of OBX as the dependent variable.....	26
Figure 1: Implied volatility (red) and realized volatility (green) over the sample period.....	15
Figure 2: OBX-VIX (red), VIX (green), and OBX (blue) over the sample period.....	16
Figure 3: Turnover on index options on Oslo Børs 2006-2020	27

1. Introduction

Investors worldwide have, throughout the times, figured out different ways of quantifying the risks related to investments. From statistical methods to complicated Greeks on equity options. The Greeks used the first options to bet on olive crops and were later used to bet on movement in equities without owning the underlying itself. Standardized option trading first started when CBOE opened its business in 1973. Recently, we have seen an increase in the use of equity options from retail traders after commission-free brokers have entered the market in the US. One of the critical aspects of valuing options is the implied volatility in the Black-Scholes model, which has proven to be a good measurement of the volatility in the equity market. The implied volatility is the unobservable expectation of the volatility of the underlying equity. The original work on a volatility index comes from Brenner and Galai in an article from 1989 (Brenner & Galai, 1989). After that, CBOE hired Robert E. Whaley to construct the Volatility Index (VIX) based on his previous work on the subject. Today, the VIX is often referred to in media as the “investor fear gauge” and measures the implied volatility on the S&P 500 index based on SPX options.

The VIX lays the foundation for our thesis, where we want to explore the characteristics of an implied volatility index based on Oslo Børs. We find it puzzling that this important measurement is not being used in more markets. We question whether this is due to liquidity or other factors making the index less accurate in other markets. Our work is inspired by Bugge et al. (2016) and their creation of the NOVIX.

This paper shows an asymmetric relationship between implied volatility and equity return in the Norwegian market. Essentially, it was necessary to construct an implied volatility index for the OBX options (OBX-VIX) to conduct our research since there is no official implied volatility index for the Norwegian market. Further, we find evidence that OBX-VIX can predict movements in the Norwegian equity market in cases where we observe the highest increases in implied volatility. We also find evidence that the asymmetric relationship is strongest in high volatility market conditions, which is in line with a previous study done in India by Chandra and Thenmozhi (2015) and the leverage effect (Black, 1976). We also find that the classic risk-reward relationship, first presented by Sharpe (1964), does not hold. Increased implied volatility lowers

investors' returns more than it generates positive returns, suggesting that the relationship is not linear. In addition, we find contradicting evidence that Simon's (2003) theory of deviations from 5-days moving average signals trends in the market.

The thesis is structured as follows; we start by presenting our research question in section 2, then we present the relevant literature related to the topic in section 3. In section 4, we will cover the methodology, and in section 5, we will present our findings.

We find the topic of implied volatility of options as a measurement for equity market volatility very exciting. We hope to shed light on an area not widely researched in Norway, as no previous studies are done using the same methodology in the Norwegian market.

2. Research question

There has been extensive research on this topic, and Chandra and Thenmozhi (2015) studied the asymmetric relationship between implied volatility and the market returns for the Indian stock market. Based on previous studies, the research question to be investigated is whether there is an asymmetric relationship between implied volatility and the underlying stock market at Oslo Børs. This is particularly interesting considering the liquidity and sector concentration in the Norwegian stock market.

Further, we want to expand the analysis to examine whether implied volatility can indicate market direction. Chandra and Thenmozhi (2015) concluded such without regressing market returns as the dependent variable. Our research is, therefore, an extension of previous studies conducted by Chandra and Thenmozhi (2015).

2.1 Testable hypothesis

Concerning our research question, we investigate the relationship between the implied volatility index for OBX and the returns from the OBX index. We start by formulating the following hypothesis:

H₀: There is not an asymmetric relationship between implied volatility and returns in the OBX Total Return Index.

H_1 : *There is an asymmetric relationship between implied volatility and returns in the OBX Total Return Index.*

We are particularly interested in investigating if specific market conditions (e.g., upward market momentum or distressed market conditions) are more applicable in providing information on the future direction of the Norwegian equity market. We will therefore expand the research with a quantile regression approach. We formulate the following hypothesis:

H_0 : *OBX-VIX does not provide explanatory information regarding the market direction.*

H_1 : *OBX-VIX does provide explanatory information regarding the market direction.*

3. Literature review

The following chapter will introduce relevant literature concerning our research question and hypothesis. The literature stems from the fundamentals of the Black-Scholes notion, key findings regarding volatility, and quantile regression.

3.1 Risk-reward relationship

Early financial research has proven a positive correlation between risk in the portfolio and expected return. Especially known is the Capital Asset Pricing Model (CAPM) developed by Sharpe (1964).

Since the risk-free rate and market risk premium are given in CAPM, the portfolio's risk is the only parameter that can be varied, implying that an increase in beta (idiosyncratic risk) will yield a higher expected return. This contradicts much of the research on the VIX index (Hibbert et al., 2008), as much of the evidence points that the VIX and returns are negatively correlated.

3.2 Black-Scholes and volatility

The market participants generally rely on past asset price movements to determine volatility expectations, looking at prior volatility of the respective asset price. Consequently, investors have a backward-looking approach to the volatility parameter. Alternatively, decision-makers can investigate implied volatility inferred from reported

option prices, providing information regarding expected volatility. The prices of options can be derived from the Black-Scholes model (Black & Scholes, 1973), where the option's value depends on the expected future volatility.

The model assumes a no-arbitrage condition, meaning that one should not be able to make risk-free profits by constructing a long-short portfolio in options and the underlying stock. Furthermore, Black and Scholes assumed ideal conditions in the market, suggesting that short-term interest rates are known and constant. The asset price follows a random walk in continuous time with a variance rate proportional to the square of the price fluctuation. The distribution of possible asset prices at the end of the interval is lognormal, and the volatility measure (variance rate of return) on the asset is constant. Also, the valuation formula assumes that the underlying stock has no kind of cash distribution or similar, and the options are of *European style*, meaning it can only be exercised at maturity (Black & Scholes, 1973).

The infamous mathematical notion for the Black-Scholes (Black & Scholes, 1973) formula:

$$C = S_t N(d_1) - K e^{-rt} N(d_2)$$

$$d_1 = \frac{\ln \frac{S_t}{K} + \left(r + \frac{\sigma^2}{2} \right) t}{\sigma \sqrt{t}}$$

$$d_2 = d_1 - \sigma \sqrt{t}$$

Equation 1: Black & Scholes option pricing model

Where:

C = Call option price

S = Current market price of underlying

K = Strike price

R = Risk-free interest rate

T = Time to maturity

N = A normal distribution

The Black-Scholes model (Black & Scholes, 1973) assumes that the underlying price follows a geometric Brownian motion with constant volatility. The original model suggests that all options with the same underlying provide the same implied volatility. However, Rubinstein (1994) examined the S&P 500 index option concluding that the

Black-Scholes model suggests that implied volatility tends to vary across strike prices and time to expiration. A similar conclusion has been drawn from examining the Philadelphia Exchange foreign currency options market by Taylor and Xu (1995). Options deep in or out of the money have the highest implied volatilities. The implied volatility tends to be lowest at-the-money, due to convexity. This relationship can be plotted into a u-shaped curve – associated with the *volatility smile*. The plotted implied volatility shape is not absolute and not perfectly formed. Near-term equity options and currency-related options are more likely to have a u-shaped curve.

3.3 Leverage effect

The leverage effect refers to the volatility and equity return relationship. More specifically, it addresses the negative correlation between asset return and volatility changes. It was first introduced by Black (1976), where the intuition is based on changes in earnings. The firm's stock price would decrease with decreasing expected earnings and increasing leverage due to increased debt to equity ratio. With the increased leverage, the volatility increases. Hence, the negative relationship between the equity returns and volatility.

3.4 Asymmetric Volatility Phenomenon

Volatility is lower in bull markets and higher in bear markets. This is known as the asymmetric volatility phenomenon (AVP). This phenomenon is well documented empirically and refers to the stylized fact that depreciation in stock prices implies higher volatility for the corresponding stock and vice-versa. (e.g. (Wu, 2001), (Bekaert & Wu, 2000)). The AVP literature has debated whether the phenomenon is more suitable for firm-specific effects such as leverage or systematic market conditions.

A study by Dennis et al. (2006) examined the relationship between implied volatility innovation, stock returns, and the asymmetric volatility phenomenon indicating that a relationship between stock returns and idiosyncratic volatility (firm-specific) is substantially near zero. Furthermore, the result suggests that asymmetric volatility is attributed to systematic volatility (market-wide).

3.5 Implied Volatility Index and stock return

Extensive studies have also examined the asymmetric relationship between implied volatility index and stock market returns. Chandra and Thenmozhi (2015) examined the relationship between the India VIX and CNX Nifty returns. The Indian implied

volatility index follows the same methodology as CBOE's white paper for constructing VIX with S&P 500 options, measuring the implied volatility for at-the-money options. Over a sample period spanning from March 1, 2009, through November 30, 2012, Chandra and Thenmozhi (2015) found that CNX Nifty returns and India VIX level is negatively correlated. Still, the relationship tends to be more independent in high upward market momentum cases. In sharp declining markets, the asymmetric relationship between implied volatility and the market becomes stronger.

3.6 Volatility clustering and mean reversion

Cont (2005) wrote a paper on volatility clustering and concluded that periods of low volatility follows low volatility and high volatility follows high volatility. He also points to the mean-reverting effects of volatility. In his Agent theory, in periods of low volatility, Agents tend to update their thresholds with small increments as they are happy with the market's pricing. Similarly, in periods of high volatility, Agents will update their threshold with large increments and become less reactive to signals, lowering the amplitude of returns. When investors get used to the new market environment, volatility will revert to its mean.

3.7 Previous work on implied volatility index in Norway

Bugge et al. (2016) created a VIX index on the Oslo Børs (NOVIX), following the methodology that CBOE used to develop the VIX for the S&P 500. Their paper also compared different indices with the NOVIX and evaluated the indices regarding volatility forecasting. They find that the NOVIX has many of the same characteristics as the VIX and the VDAX-NEW (volatility index for the German market). The negative correlation between OBX and NOVIX became stronger during the sample period and is closer to the correlation of S&P 500 VIX and VDAX-NEW in 2016. This makes us more confident in using Norwegian equity volatility to determine if it can be used to forecast equity returns.

3.8 Quantile regression approach

The highly cited research from Koenker and Hallock (2001) is an extension of the quantile regression first introduced by Koenker and Bassett (1982). The concept sheds light on conditional quantile functions where quantile regression models the relationship between independent variables and specific quantiles of the dependent variables. Furthermore, the approach models the relationship to a target, most often set

to the median of the distribution. Quantile regression makes no assumptions about the distribution of the dependent variable, which is different from ordinary least squares regression. Quantile regression tends not to be influenced by outlying observations and is applicable to regress OBX-VIX and OBX returns.

4. The research methodology and design

This chapter will go through the research design and methodologies used in our analysis of OBX-VIX. All computations and analyses are done in Python.

4.1 Research design

Our research design for our research question can be divided into four stages. In stage 1, we collect historical data for the OBX option from Oslo Børs to construct our implied volatility index, OBX-VIX. We start by identifying the relevant options according to CBOE's white paper and calculate the implied volatility for near-term and next-term options.

In stage 2, we test the robustness of OBX-VIX, where we conduct descriptive analysis and calculate correlation to compare the findings to the well-established VIX index from CBOE.

Stage 3 involves running a multiple regression analysis of the OBX-VIX concerning key variables addressed in the methodology.

In stage 4, we run a quantile regression analysis of OBX-VIX and OBX returns to test the relationship between implied volatility and OBX returns.

4.2 Methodology

This section will examine the methods used in our research and why they were chosen. We will explain how we created OBX-VIX, one of the key input variables in our different regression models.

4.2.1 VIX construction

In creating OBX-VIX, we have opted to use the same methodology as CBOE in creating the VIX. This is the industry standard way of calculating a model-free implied volatility approach on underlying indices. We believe this will give us the most

accurate results as the methodology has been tested in multiple markets. In this section, we will present critical calculations used in the creation.

$$\sigma^2 = \frac{2}{T} \sum_i \left(\frac{\Delta K_i}{K_i^2} e^{RT} Q(K_i) - \frac{1}{T} \left[\frac{F}{K_0} - 1 \right]^2 \right)$$

Equation 2: The general formula to calculate VIX

Where:

$\sigma = \text{VIX}/100 \rightarrow \text{VIX} = \sigma \times 100$

T = Time to expiration

F = Forward index level derived from index option prices

K_0 = First strike below the forward index level, F

K_i = Strike price of i^{th} out-of-the-money option; a call if $K_i > K_0$ and a put if $K_i < K_0$; both put and call if $K_i = K_0$

R = Risk-free interest rate to expiration

$Q(K_i)$ = The midpoint of the bid-ask spread for each option with strike K_i

ΔK_i = Interval between strike prices – half the difference between the strike on either side of K_i

$$\Delta K_i = \frac{K_{i+2} - K_{i-1}}{2}$$

Equation 3: Interval between strike prices in VIX calculation

In the OBX-VIX, we will use near- and next-term put and call options with more than 23 days to expiration and less than 37 days to expiration. In the case of a lack of option quotes, we have made an exception to the rule in our code. If there are no options with 30 days to maturity, we allow the next best option to be chosen. We measure TTM in minutes to have the precision most often used in options trading.

$$T = \frac{\{M_{\text{Current day}} + M_{\text{Settlement day}} + M_{\text{Other days}}\}}{\text{Minutes in a year}}$$

Equation 4: Time to maturity in VIX calculation

Where:

$M_{\text{Current day}}$ = Minutes remaining until midnight of the current day

$M_{\text{Settlement day}}$ = Minutes from midnight until 16:30 for “standard” OBX expirations

$M_{\text{Other days}}$ = Total minutes in the days between the current day and expiration day

The risk-free rates are based on the Norwegian Treasury yields at the options expiration. We select out-of-the-money OBX calls for the options used, and out-of-the-money OBX puts with non-zero bids, close to the strike price, K_0 . To find F , we calculate the difference between calls and puts to find the strike where the difference is smallest. We do this for both near- and next-term options. Next, we determine K_0 , the strike price equal to or immediately below the forward index level, F . After that, we select out-of-the-money put options with strike prices $< K_0$. We work our way down, exclude any options with no bid, and stop when we reach two successive puts with no bids. We perform the same calculation upwards for the call options with strikes $> K_0$. Lastly, we find the call and put with strike = K_0 and the midpoint price, using the average.

$$F = \text{Strike Price} + e^{RT} * (\text{Call Price} - \text{Put Price})$$

Equation 5: Future price on underlying asset in VIX calculation

Calculate volatility for both near-term and next-term options:

$$\sigma_1^2 = \frac{2}{T_1} \sum_i \frac{\Delta K_i}{K_i^2} e^{R_1 T_1} Q(K_i) - \frac{1}{T_1} \left[\frac{F_1}{K_0} - 1 \right]^2$$

$$\sigma_2^2 = \frac{2}{T_2} \sum_i \frac{\Delta K_i}{K_i^2} e^{R_2 T_2} Q(K_i) - \frac{1}{T_2} \left[\frac{F_2}{K_0} - 1 \right]^2$$

Equation 6: Volatility for near- and next-term options in VIX calculation

Each option added to the index is proportional to ΔK and the price. It is inversely proportional to the square of the option's strike. In general, ΔK will be half the difference between the strike prices on either side of K_i .

To find the 30-day weighted average of σ_1^2 and σ_2^2 :

$$OBX - VIX = 100 * \sqrt{\left\{ \left[\frac{N_{T_2} - N_{30}}{N_{T_2} - N_{T_1}} \right] + T_2 \sigma_2^2 \left[\frac{N_{30} - N_{T_1}}{N_{T_2} - N_{T_1}} \right] \right\} * \frac{N_{365}}{N_{30}}}$$

Equation 7: 30-day weighted average value of the volatility in VIX calculation

Where:

N_{T_1} = number of minutes to settlement of the near-term options

N_{T_2} = number of minutes to settlement of the next-term options

N_{30} = number of minutes in 30 days

N_{365} = number of minutes in a 365-day year

Using monthly options will result in the index always reflecting an interpolation of σ^2_1 and σ^2_2 . For our construction based on the Norwegian market, we have made some minor alterations as a perfect replication of the methodology used by CBOE is not feasible with our data. To ensure that we get credible readings, we have had to include options with maturities widening outside the 24 to 36 days window used by CBOE. The complete code can be found in the appendix.

The liquidity of the Norwegian option market made it more difficult to calculate an accurate representation of the implied volatility. This led to us not having daily data for the entire sample period but some weeks with three or four data points. We also had to exclude some of the outliers on the low side of the index due to obvious calculation errors because of the low liquidity.

For in-depth information on the construction, we refer to the CBOE white paper (*Cboe VIX White Paper*, n.d.).

4.2.2 Multiple Linear Regression

Denote the observed log prices with daily data from Oslo Børs and the implied volatility output from the OBX-VIX index of the Norwegian equity market; we follow a regression-based methodology to examine any asymmetric relationship between the indices. We incorporate findings from Simon (2003) suggesting that technical indicators might signal market trends. More significant positive deviations of OBX returns from its 5-day MA (moving average) may indicate market sentiment. We will analyze the empirical features of implied volatility established from OBX and the equity return by conducting the following multiple regression:

$$\Delta OBXVIX_t = \beta_0 + \beta_1 OBXVIX_{t-1} + \beta_2 OBXret_t^+ + \beta_3 OBXret_t^- + \beta_4 DevMA5_t^+ + \beta_5 DevMA5_t^-$$

Equation 8: Multiple Linear Regression model

Where:

$\Delta OBXVIX_t$ = first difference of OBX-VIX in t

$OBXVIX_{t-1}$ = the lagged value of OBX-VIX

$OBXret_t^+$ = positive OBX return in t

$OBXret_t^-$ = negative OBX return in t

$DevMA5_t^+$ = positive deviation term of 5-day moving average from the closing price of OBX in t

$DevMA5_t^-$ = negative deviation of 5-day moving average from the closing price of OBX in t

By the above multiple regression, we will study the statistical properties of OBX-VIX and the underlying equity market by looking at the dependent variable $\Delta OBXVIX_t$ (first difference) in respect to the performance of OBX returns. By separating positive and negative OBX returns and including deviation from the 5-day MA in the multiple linear regression model, we hope to capture the relationship dynamics between the implied volatility index and the underlying index that can indicate future stock market performance.

According to Simon (2003), deviation from the recent market tendency (5-day MA) can indicate market trends and the respective trend's direction. He argues that market participants tend to demand more call options in upward market conditions (above 5-day MA), which results in investors bidding up the implied volatility. In such situations, the positive (negative) deviation coefficient would be significantly positive (negative). Therefore, including the deviation terms will help us analyze the anticipated asymmetric relationship between OBX-VIX and the equity market changes with respect to deviations from the stock market tendency.

4.2.3 Quantile Regression

We assume the financial time series has a leptokurtic distribution and will therefore examine the statistical properties of OBX-VIX and the OBX returns using a quantile regression approach to capture a more comprehensive view of the relationship of the data. First, the regression technique estimates the conditional median function, and then

the entire range of the other conditional quantile functions. This allows us to obtain a picture of the behavior of returns and implied volatility in different quantiles. Koenker and Hallock (2001) found that quantile regression better measures outliers in the time series.

First, we establish the quantile regression with the following for the τ -th conditional quantile function of OBX-VIX. The parameter $y(\tau)$ is the effect of OBX returns at the respective quantile of the conditional distribution.

$$QOBXVIX_t(\tau|OBXret_{t-1}, x_t) = y(\tau) OBXret_{t-1} + x'_t(\tau)$$

Equation 9: Conditional quantile function

Second, to estimate the model above, we solve it for $\rho_t(u)$, which is the standard quantile regression check function ((Koenker & Bassett, 1982), (Koenker, 2005)). The result from the equation will be set as the pooled quantile regression estimator.

$$\min_{\gamma, \beta \in \mathcal{C} \times \mathcal{B}} \sum_{t=1}^T \rho_t(OBXVIX_t - y(OBXret_{t-1} - x'_t \beta))$$

Equation 10: Minimizing problem equation

Lastly, we examine the relationship between OBX returns and the implied volatility index performance by estimating the quantile regression model from equation 11. The quantiles we choose to investigate range from 0.1 to 0.9 to get a better picture of any outliers. We have selected these quantiles to minimize the probability that different market conditions (high implied volatility and low implied volatility market conditions) get mixed in the same quantiles. We expect that the lowest and highest quantiles will provide the most robust results from the regression.

$$OBXVIXret_t = \alpha_0 + \alpha_1 OBXret_t^+ + \alpha_2 OBXret_t^-$$

Equation 11: Quantile regression model

Where:

$OBXVIXret_t$ = return of OBX-VIX in time t

α = respective coefficients

4.3 Data collection

In this section, we introduce the method for data collection, including sources used to collect data. Moreover, we will present necessary data handling for constructing the OBX-VIX.

4.3.1 OBX option

The BI Norwegian Business School Library provided the raw OBX option data files in arrangement with Oslo Børs. The data consists of daily closing bid and ask quotes, strike prices, and days to expiry for all put and call options and was initially stored in the OBI database. However, the dataset was incomplete, with missing data for 2009 and 2020.

To solve this issue, we substituted the missing data with an earlier version of an implied volatility index (NOVIX) which follows the same methodology for calculating the implied volatility index, using the OBX Total Return Index options.

The following data has been used to construct the implied volatility index OBX-VIX:

Data	Source	Data Range
OBX options	OBI database	2007.01.02-2008.12.30
NOVIX Index	Bugge et al.	2009.01.02-2009.12.30
OBX options	OBI database	2010.01.04-2019.12.30
NOVIX Index	Bugge et al.	2020.01.01-2020.05.22

Table 1: Overview of data used to construct the OBX-VIX index.

The data starts in 2007 to capture the financial crisis. The sample period will include several periods of market distress and upward momentum to capture any outliers in the financial time series. We can better picture the relationship between the OBX return and OBX-VIX in these events. Furthermore, the sample period ends in May 2020 since Euronext acquired Oslo Børs, and the stock exchange had a transition to new data reporting and storage systems. All other data in our thesis corresponds with the time series described above in Table 1.

According to Anders Holen, some market makers on Oslo Børs tend to pull screen prices right before market close, meaning that the data set, including bid and ask quotes at market close, does not necessarily give the correct picture of the market. This may

explain some of the missing data points but could be mitigated by choosing a different time during the day to collect the data. This, however, requires high-frequency data.

4.3.3 OBX Total Return index and Risk-free rate

We have collected daily closing prices for the OBX Total Return Index from Refinitive Eikon. The index currently constitutes 82% of the market capitalization of the OSEBX index and is an indicator of the overall performance of the Oslo Stock Exchange.

We used the 12 months Norwegian Treasury yield curve for the risk-free rate. The data was fetched from Norges Bank's API for interest rates.

5. Results

In this chapter, we will present our results based on our analysis. We will go into detail on correlations and our different regression models and summarize the findings giving an economic interpretation of our findings.

5.1 Descriptive analysis

Table 2 summarizes descriptive statistics for OBX-VIX and VIX in level terms and returns. In OBX-VIX, we can see that the maximum value is 74.92, observed 2008-10-16. The maximum value occurred at the break-out of the financial crisis in 2008, where we also saw the VIX reach its maximum value of 82.69. The table shows that the mean of the returns of both volatility indices is close to zero but slightly positive. The data has an excess kurtosis above 0, implying that the data is leptokurtic, with fatter tails than the normal distribution, in line with most financial data. The skewness of the data is closer to zero on returns, but the absolute values of the indices have a positive skewness, which might imply that the data is slightly right-skewed. In figure 1, we have plotted the implied volatility and the realized volatility over the sample period to visualize the data.

	Mean (%)	Std. (%)	Min. (%)	Max. (%)	Skew.	Ex. Kurt.
OBX-VIX	22.43	9.31	10.03	74.92	1.97	5.21
OBX-VIX ret	0.01	5.78	-50.49	43.51	0.20	9.35
VIX	19.68	9.83	9.14	82.69	2.48	8.03
VIX ret	0.03	7.78	-35.06	76.83	1.04	6.12

Table 2: Descriptive statistics of OBX-VIX and VIX and their respective returns.

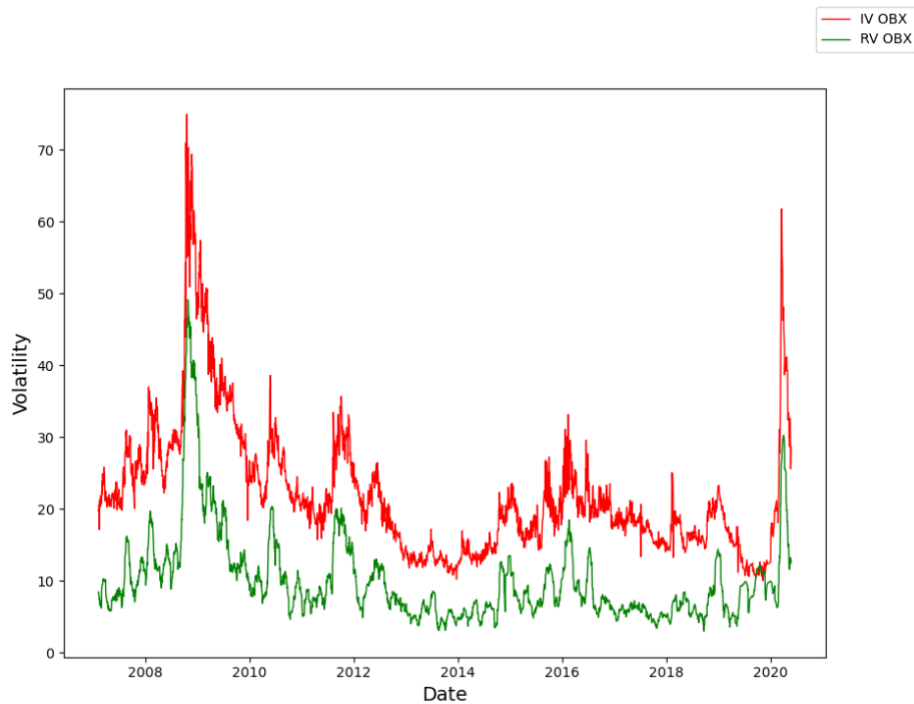


Figure 1: Implied volatility (red) and realized volatility (green) over the sample period.

5.2 Correlation

As presented in table 3, we tested different variables in different sample periods. We can see the negative correlation between OBX-VIX and OBX becomes stronger, from -0.340 in 2007-2010 to -0.456 in the last sample, and the correlation is approaching the results from S&P 500 and VIX in the same sample period. We believe that looking towards the correlation between S&P 500 and VIX as a benchmark value since the US market is the most liquid in the world. Our results are similar to those of

Bugge et al. (2016). In figure 2, we have plotted the time series together to visualize the findings.

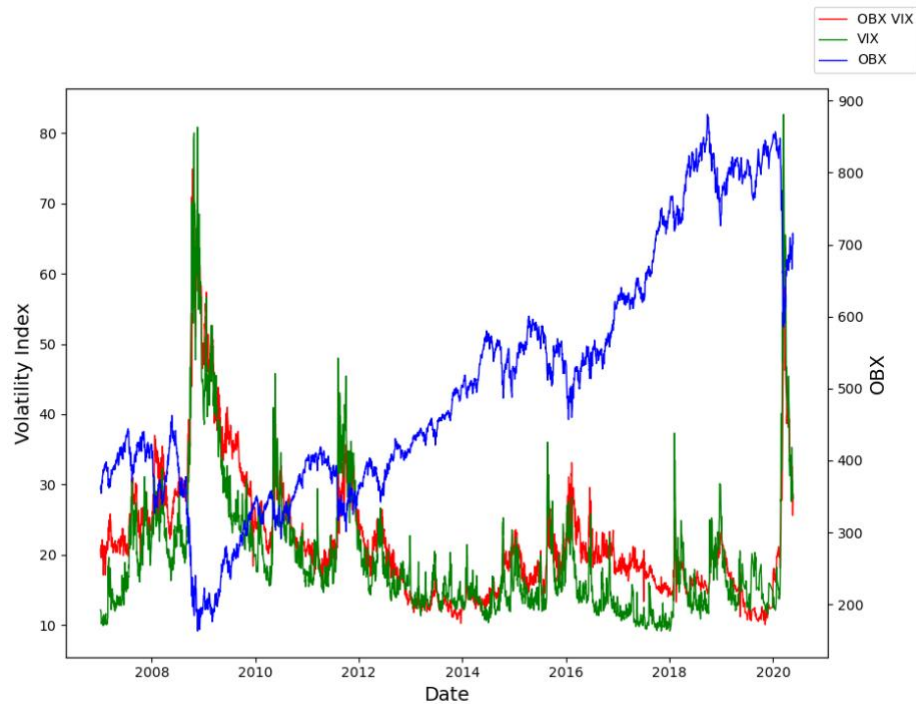


Figure 2: OBX-VIX (red), VIX (green), and OBX (blue) over the sample period.

The more surprising result is how correlated the volatility indices are. Intuitively, one would believe that there would be more significant discrepancies in what risk factors affect investors when investing in different markets. Even though the economy and financial markets are globalized, we still believe that the difference in industry weighting in the underlying indices would be exposed to different risks and macroeconomically factors that affect the markets differently. The strong correlation between the two volatility indices begs the question if the creation of an official OBX-VIX which will enable trading volatility is necessary. With the already low liquidity in the Norwegian derivative market, it would be hard to justify the creation of this index due to potential high liquidity costs for investors.

In addition, we calculated the 30-day realized volatility on OBX and S&P 500. We wanted to test how well the implied volatility could be used to predict the realized volatility. The phenomenon is well researched; for example, Jiang & Tian (2006) found that VIX subsumes all information contained in the realized volatility in the S&P 500. After analyzing the correlation between RV OBX and OBX-VIX, we find similar

results, indicating that implied volatility is also a suitable method of forecasting realized volatility.

	Entire sample 2007-2020	First period 2007-2010	Second period 2011-2015	Third period 2015-2020
OBX-VIX and VIX	0.906	0.948	0.844	0.859
RV** OBX and RV S&P 500	0.895	0.957	0.829	0.877
OBX and S&P 500	0.538	0.533	0.611	0.499
S&P 500 and VIX	-0.731	-0.762	-0.827	-0.688
OBX and OBX-VIX	-0.363	-0.340	-0.401	-0.456
OBX and VIX	-0.394	-0.420	-0.462	-0.373
RV OBX and OBX-VIX	0.906	0.924	0.857	0.811

Table 3: Correlation between the different variables across different sample periods.

* All results are statistically significant at the 1% level.

** 30-days realized volatility calculated as:

$$r_i = \ln(p_t) - \ln(p_{t-1})$$

$$RV = \sum_{i=1}^T r_i^2 * \sqrt{252}, \text{ with } r_i^2 \text{ being calculated with a 30-day rolling window}$$

5.3 Multiple Linear Regression

5.3.1 Model and results

In our analysis of equation 12, we have separated the regression into six models (table 4). This technique has been used to measure better the degree to which independent variables provide explanatory power to the dependent variable separately. We also want to study if expanding the regression with independent variables will have a meaningful impact on the adjusted R-square and regressed coefficients. In other words, we will be able to evaluate which model is best suited to comprehend the relationship between implied volatility and the performance of the underlying equity market. The

regression for all models has used the entire sample period, where table 4 shows the results from the multiple regression for the respective models.

In our regression, we study how the lagged value of OBX-VIX ($OBXVIX_{t-1}$) has any explanatory power to the change of OBX-VIX ($\Delta OBXVIX_t$), together with the performance of the underlying equity market. We have separated positive ($OBXret_t^+$) and negative ($OBXret_t^-$) OBX returns to identify better the expected asymmetric relationship between implied volatility and index performance. In addition, we have added positive ($DevMA5_t^+$) and negative ($DevMA5_t^-$) 5-day moving average deviation in percentage to the regression to capture any market trends.

The multiple regression is based on equation 13:

$$\Delta OBXVIX_t = \beta_0 + \beta_1 OBXVIX_{t-1} + \beta_2 OBXret_t^+ + \beta_3 OBXret_t^- + \beta_4 DevMA5_t^+ + \beta_5 DevMA5_t^-$$

Equation 12: Multiple Linear Regression with the first difference of OBX-VIX as the dependent variable

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Constant	0.0032 (0.026) [0.124]	0.2383 (0.029) [8.250]***	-0.2692 (0.027) [- 10.031]***	-0.0617 (0.031) [-2.004]**	0.00733 (0.032) [0.226]	-0.0115 (0.033) [-0.345]
Lagged value of OBX VIX	-6.2286 (0.448) [- 13.908]***	-5.9545 (0.431) [- 13.805]***	-6.3864 (0.416) [- 15.359]***	-7.6958 (0.414) [- 18.586]***	-7.3530 (0.408) [- 18.013]***	-7.9423 (0.422) [-18.826]***
Positive OBX returns		-42.4658 (2.648) [- 16.035]***		-31.4949 (2.545) [- 12.375]***		-11.4422 (3.218) [-3.555]***
Negative OBX returns			-51.2293 (2.255) [- 22.717]***		-42.2740 (2.269) [- 18.628]***	-20.7508 (3.455) [-6.006]***
Positive Dev from MA5				40.9188 (1.984) [20.626]***		23.9073 (3.022) [-7.910]***
Negative Dev from MA5					38.0467 (2.269) [- 14.558]***	28.6111 (3.271) [-8.746]***
R square	0.057	0.127	0.187	0.229	0.238	0.254
Adjusted R square	0.057	0.126	0.187	0.228	0.237	0.253
F-statistic	193.4	233.0	370.3	317.7	333.7	218.3

Table 4: Multiple Linear Regression, a summary of statistics from the six different models.

SE and t-statistics are reported in () and [] respectively.

*** Significant at 1% level; ** Significant at 5% level.

Model 1 estimates the coefficient of the lagged value of OBX-VIX with respect to the change in OBX-VIX. From the regression, the constant variable is insignificant at all levels. We get a significant coefficient of -6.2286 with a 1% significance level. This implies that the change in OBX-VIX at time t will be negatively impacted by the absolute value of OBX-VIX the previous day.

In model 2, we include positive OBX returns ($OBXret_t^+$) to the regression. The lagged value of OBX-VIX and positive OBX returns are negatively related to the change in OBX-VIX and significant at a 1% significance level. The constant is also significant at the same significance level. Moreover, including positive OBX returns as an independent variable has increased the adjusted R-square to 0.126.

Model 3 examines the explanatory power of positive OBX returns and the lagged OBX-VIX regarding the change in OBX-VIX. All coefficients and the constant term are significant at a 1% significance level and negatively related to the change in OBX-VIX. Furthermore, by replacing the positive OBX return with negative OBX returns as an independent variable, the adjusted R-squared has increased from 0.126 to 0.187. The coefficient from negative OBX returns is also greater than in model 2. Thus, negative returns in the underlying index have a greater statistical and economic explanatory power to the implied volatility change than positive returns. This finding is according to our expectations as investors are more concerned about the *downside risk* than *upside risk* and are in line with the results of Kahneman and Tversky (1979).

Model 4 considers the lagged value of OBX-VIX, positive OBX returns, and the deviation term capturing the positive 5-day moving average digression from the market tendency. The constant is significant at a 5% significance level from the regression, while all other coefficients are significant at a 1% significance level with an adjusted R-square equal to 0.228. The independent variables lagged OBX, and positive OBX returns remain negatively related to the change in OBX-VIX. The coefficient from the positive 5MA deviation term is positively related to the change in OBX-VIX. Our results reveal that a positive deviation in percentage from the current market trend positively impacts the change in OBX-VIX. In other words, the implied volatility increases when the market trend is above its recent 5-day tendency.

Model 5 estimates the statistical and economic explanatory power from the independent variables lagged OBX-VIX, negative OBX returns, and the negative 5MA deviation term in relation to the changes in OBX-VIX. The constant of model 5 is not significant. Meanwhile, the remaining coefficients are significant at all significance levels. The lagged OBX-VIX variable and negative OBX returns remain negatively related to the change in OBX-VIX. Interestingly, we observe that the negative 5MA deviation term is positively related to the change in OBX-VIX and differs from the findings of Simon (2003). Economically, a market trend below the recent 5-day tendency decreases the implied volatility. We expected a negative coefficient because investors would bid up implied volatility by demanding put options in a downward trend. However, the findings do not necessarily provide sufficient economic explanation since the mean of negative 5-day moving average deviation in percentage from the market equals negative 0.6%. Thus, it is suggested that on average that the 5MA deviation term will affect the change in OBX-VIX only equal to 0.228 (absolute term). The adjusted R-square from model 5 equals 0.237.

Model 6 considers the lagged value of OBX-VIX, positive OBX returns, negative OBX returns, positive 5MA deviation, and negative 5MA deviation in relation to the change in OBX-VIX. All estimated coefficients are significant at a 1% significance level without the constant term, which is not significant. The adjusted R-square is the highest in model 6, equaling 0.253, and therefore has the highest explanatory power of all the models.

Furthermore, the lagged value of OBX-VIX being negatively related to the change in OBX-VIX suggest that implied volatility is mean-reverting. This is because volatility reverts to its mean, according to Cont (2005). Our results also show that positive and negative OBX returns are negatively related to OBX-VIX. Higher positive OBX returns will decrease the OBX-VIX. A 1% increase in the index results in an 11.44% decrease in OBX-VIX, and a 1% decrease in the OBX index will increase the OBX-VIX by 20.75%. This is consistent with the findings from the other models, where investors are more sensitive to declining markets than increasing markets. Economically, the negative OBX return has a greater impact because put options generally have a higher premium than call options, suggesting higher expected volatility in the downfall.

Regarding the positive and negative 5MA deviation terms, we find them contradictory to the findings from Simon (2003). In model 6, both coefficients were significant at 1%. Still, the coefficient for negative deviation is positively related to changes in OBX-VIX, suggesting that when the market is below its 5-day tendency, the change in OBX-VIX decreases. This might be the case since market makers at Oslo Børs, on average, anticipate lower implied volatility in these conditions, or the demand for put options is poor where we observe minor deviations from the 5-day tendency, reducing their option quotes. However, using the 5MA deviation term has low economic explanatory power since the average positive, and negative deviation from its 5MA equals respectively 0.6% and -0.6% resulting in a low impact on the change of OBX-VIX.

From our overall results from the multiple linear regression, it is evident that positive returns in the market decrease the implied volatility because this reduces the negative market sentiment, and negative market returns increase the implied volatility, where the latter has the most significant impact. Moreover, we see a tendency from our results that an increase (decrease) in the implied volatility index is subsequent to a decrease (increase) in OBX returns for the same trading day.

5.3.2 Linear regression assumptions and our data

There are some complications in using financial time-series data concerning the classic linear regression assumptions that lay the foundation for the multiple linear regression. This is especially true for the homoscedastic and normally distributed error terms assumptions. By performing White's test and Jarque-Bera's tests on our models, we can see that the data violates two of the classical assumptions. On the other hand, we have such a large sample size that we believe the effects on our results will be minimal.

5.4 Quantile Regression: OBX-VIX as the dependent variable

The key observation from our multiple regression analysis was the asymmetric relationship between the implied volatility index and the underlying equity market. We expand our research by conducting a quantile regression analysis to get a more comprehensive picture of this relationship.

In the case of quantile regression, we seek to estimate the expected asymmetric relationship between implied volatility and returns, studying quantiles of the

conditional distribution of the response variable (Koenker & Hallock, 2001). The quantile function for each quantile can be addressed as follows:

$$QOBXVIX_t(\tau|OBXret_{t-1}, x_t) = y(\tau) OBXret_{t-1} + x'_t(\tau)$$

Equation 13: Quantile function for each quantile

The parameter captures the impact OBX returns contribute to at the (τ) -quantile of the conditional distribution of the OBX-VIX index. Following the methodology described in 4.2.3, we pool the quantile regression estimators by solving equation 10.

We obtain our quantile regression analysis results by regressing the following equation. The results from the quantile regression are gathered in table 5.

$$OBXVIXret_t = \alpha_0 + \alpha_1 OBXret_t^+ + \alpha_2 OBXret_t^-$$

Equation 14: Quantile regression with returns of OBX-VIX as the dependent variable

Quantile	Constant			OBX ret (+)			OBX ret (-)		
	Coefficient	SE	t-Statistic	Coefficient	SE	t-Statistic	Coefficient	SE	t-Statistic
0.100	-0.0337	0.002	-16.327***	-2.442	0.145	-15.499***	-0.1898	0.175	-1.086
0.200	-0.0168	0.001	-17.076***	-2.1646	0.070	-30.918***	-0.5981	0.080	-7.515**
0.300	-0.0110	0.001	-13.764***	-1.8167	0.060	-30.306***	-0.9755	0.063	-15.499***
0.400	-0.0068	0.001	-9.058**	-1.5138	0.059	-25.492***	-1.2418	0.056	-22.086***
0.500	-0.0027	0.001	-3.437**	-1.3113	0.065	-20.136***	-1.5727	0.058	-27.339***
0.600	0.0020	0.001	2.294**	-1.0226	0.075	-13.678***	-1.8800	0.061	-30.890***
0.700	0.0085	0.001	8.631**	-0.8838	0.087	-10.145***	-2.1439	0.066	-32.441***
0.800	0.0174	0.001	12.632***	-0.7083	0.127	-5.575**	-2.4289	0.087	-27.908***
0.900	0.0335	0.002	15.014***	-0.3510	0.218	-1.613	-3.0502	0.132	-23.141***

Table 5: Quantile regression, summarizing statistics from quantile 0.1-0.9.

*** Significant at 5% level; ** Significant at 1% level.

All constant terms are statistically significant from the quantile regression with a 1% significance level. This violates the stylized facts of volatility being mean-reverting since we can display a significant negative trend from the constant term from all quantiles from the regression. In the case of mean-reverting volatility, we should not be able to observe any significant trend. This finding is according to evidence provided by Chandra and Thenmozhi (2015) from their quantile regression for the Indian market and Siriopoulos and Fasas (2008) for multiple markets.

The estimated coefficients are statistically significant at 1% for all quantiles except the lower quantile (0.1) for negative OBX returns and upper quantile (0.9) for positive OBX returns, where the coefficients are insignificant. The results suggest that the relationship between OBX returns and OBX-VIX returns are negatively related in both directions. This is consistent with our findings from the multiple regression analysis.

As expected, the results hold more for market declines than upward market conditions. It becomes evident that the negative relationship between negative OBX returns and OBX-VIX becomes stronger in the quantiles where the volatility index increases the most. Economically this can be interpreted as investors' expectations of future volatility to increase more in a market downturn, accordingly to volatility clustering (Cont, 2005).

We see that positive OBX returns have a low estimated coefficient in the upper bound of the quantiles. Therefore, the market will not be comforted with positive returns when the implied volatility is already high. However, there is a significant negative relationship for both positive and negative OBX returns, and it is evident that substantial leverage effects are present.

5.5 Quantile Regression: OBX return as the dependent variable

The results from the quantile regression above also imply that returns in the underlying market are related to the returns in the implied volatility index. Looking at this the other way, the returns in OBX-VIX contribute to the movements in OBX returns because it changes the market sentiment and market participants position themselves with options instead of the underlying equity, causing delta hedging from the market maker.

However, this relationship is vaguer and will be further analyzed by reversing the quantile regression from section 5.4.

By pooling the quantiles with respect to the return of the OBX-VIX index, we can examine whether there is a negative relationship between the OBX return (dependent variable) and lagged positive and negative OBX-VIX returns (independent variables) in different implied volatility environments. The highest quantile has the highest increases in OBX-VIX, and the lowest quantile includes market situations where the return in OBX-VIX is the lowest. The quantiles go from 0.1 to 1.

We obtain our results from our quantile regression analysis by regressing the following equation:

$$OBXret_t = \alpha_0 + \alpha_1 OBXVIXret_{t-1}^+ + \alpha_2 OBXVIXret_{t-1}^-$$

Equation 15: Quantile regression with returns of OBX as the dependent variable

From the quantile regression, we observe that only in the highest quantile can we obtain significant results at a 1% significance level. At this quantile, the return on the OBX-VIX index is the highest, meaning the anticipated volatility has increased. The estimated coefficient for positive OBX-VIX returns equals -0.1473, while negative OBX-VIX returns remain insignificant. The constant term is significant at a 1% significance level with a coefficient of -0.0128. This implies that increased implied volatility will negatively impact the market performance the following day in highly distressed market conditions.

From the results, we see that in highly distressed market conditions, the selling pressure will increase with a hike in the implied volatility index. This could often be a reallocation from risky assets since investors are more sensitive to downside risk. The reason for the estimated coefficient for negative OBX-VIX returns being insignificant is because the distribution within the highest quantile only consists of positive returns in the OBX-VIX. At this quantile, the interval includes OBX-VIX returns from 5.63% to 43.5%, meaning with such changes in the implied volatility index, we can anticipate a negative impact on the market performance the following day.

5.6 Discussion and findings of empirical evidence

This section will summarize our findings from the previous chapters and discuss the economic implications of these results. We started with a descriptive analysis of the data and investigated the correlation between variables. We found that the descriptive statistics between OBX-VIX and VIX were very similar, giving us a quick and good indication that we had successfully created a valid volatility index in the Norwegian market. This was further confirmed by our results in the correlation, as we saw that the correlation of the OBX-VIX and OBX closed in on the correlation between VIX and S&P 500. The more interesting point here is the reason for the stronger correlation over the sample period. We believed the stronger correlation was due to more activity in the options market, but as figure 3 shows, there has been a steady decline in turnover in the index options market on Oslo Børs. After the crash in 2008, most investment banks stopped trading options to take on risk, which had been very popular up until this point. With new regulations regarding this in the US, the effect hit the Norwegian market. There must therefore be another explanation for the stronger correlation. A possible reason could be that technological advancement could provide market makers with more correct pricing of volatility across markets, as market makers often operate in multiple markets globally.

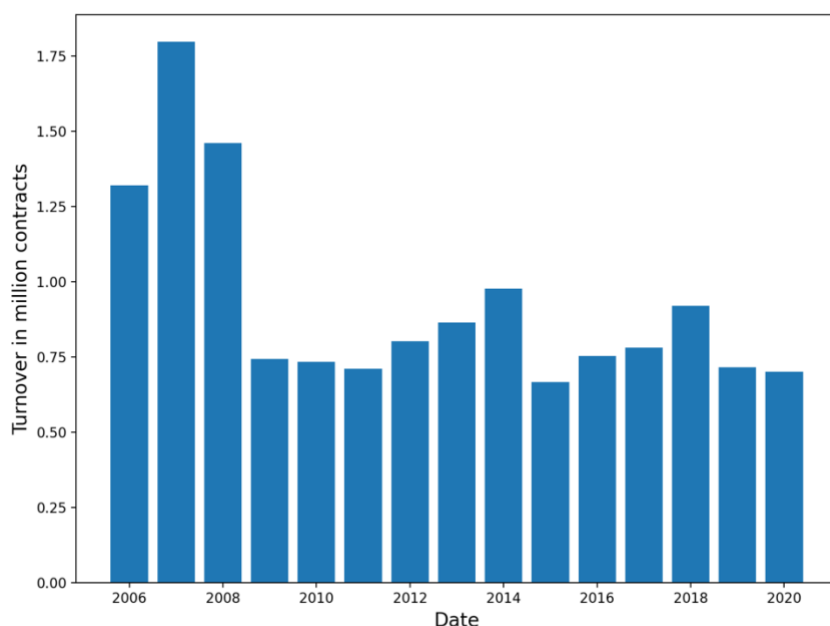


Figure 3: Turnover on index options on Oslo Børs 2006-2020

Regarding our hypotheses, we believe we have found sufficient evidence to reject our first null hypothesis and conclude that we have discovered an asymmetric relationship between implied volatility and returns on Oslo Børs. The findings are more apparent in market downturns than upward moving markets, which the leverage effect and the asymmetric volatility phenomenon can partly explain. Regarding our second hypothesis, we find that the relationship is apparent both ways and we can reject the second null hypothesis as well. Extreme returns in OBX-VIX will negatively affect OBX returns the following day.

The asymmetric relationship between the stock market returns and OBX-VIX demonstrates that the expected behavior of implied volatility in the Norwegian equity market is according to stylized facts and established theory. However, from our multiple regression analysis, the theory from Simon (2003) does not entirely hold since the coefficient for the negative deviation term ($DevMA5_{-1}$) was positive. We expected this coefficient to be negative, suggesting that implied volatility increases with a percentage deviation from the recent market tendency because investors prefer put options in such cases, driving implied volatility upwards.

We have also used quantile regression to examine whether this asymmetric relationship is stronger in different market conditions. This was the case for sharp market declines and strong upward market momentum for respectively negative and positive market returns. Interestingly, we found a violation of implied volatility being mean-reverting for the quantiles from the quantile regression. This is, however, not necessarily the case. The quantiles obtained less than 400 data points, and we found strong evidence of the implied volatility being mean-reverting from the multiple regression with the entire sample.

We recommend a larger sample size spanning more years for future research. Our sample has been from a period spanning from one of the most significant market crashes in modern times to the longest bull-run we have seen, complemented by two other significant crises. We believe the evidence would be more apparent and more rooted in normal market conditions with an extended period of data. We would also recommend utilizing high-frequency data due to new information being priced in by

market participants in seconds rather than days. In addition, this will prevent the data issue caused by market makers pulling prices on close.

When we started with the thesis, we also wanted to see what value OBX-VIX could bring to the market and if it would be feasible with an official version of the index at some point. From our research, we can conclude that this will probably not be worthwhile for other than research purposes. As seen in the chapter on correlation, OBX-VIX and VIX have a strong correlation over the sample period meaning that investors will get close to the same exposure trading derivatives on both indices. As discussed with Anders Holen, Norway's options market has been relatively slow in the last few years. There does not seem to be a market for providing derivatives on volatility.

6. Conclusion

In this paper, we have constructed an implied volatility index for Oslo Børs. We have looked at characteristics, compared them to VIX, and looked at asymmetric relationships between implied volatility and returns in the Norwegian market. We found that the relationship is more evident in market downturns than upward market tendencies. In addition, we have researched if the relationship can be found both ways. There is clear evidence that returns affect implied volatility and the other way around.

Bibliography

Bekaert, G., & Wu, G. (2000). Asymmetric Volatility and Risk in Equity Markets.

The Review of Financial Studies, 13(1), 1–42.

<https://doi.org/10.1093/rfs/13.1.1>

Black, F. (1976). Studies of Stock Price Volatility Changes. *Proceeding of the 1976*

Meetings of the American Statistical Association, Business and Economic

Statistics Section, 177–181.

Black, F., & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities.

Journal of Political Economy, 81(3), 637–654. JSTOR.

Brenner & Galai. (1989). New Financial Instruments For Hedging Changes In

Volatility. *Financial Analysts Journal*, 45(4), 61.

Bugge, S., Guttormsen, H., Molnár, P., & Ringdal, M. (2016). *Implied Volatility*

Index for the Norwegian Equity Market (SSRN Scholarly Paper ID 2816799).

Social Science Research Network. <https://papers.ssrn.com/abstract=2816799>

Cboe VIX White Paper. (n.d.). Retrieved September 12, 2021, from

<https://cdn.cboe.com/resources/vix/vixwhite.pdf>

Chandra, A., & Thenmozhi, M. (2015). On asymmetric relationship of India volatility

index (India VIX) with stock market return and risk management. *DECISION*,

42(1), 33–55. <https://doi.org/10.1007/s40622-014-0070-0>

Cont, R. (2005). Volatility Clustering in Financial Markets: Empirical Facts and

Agent-Based Models. *SSRN Electronic Journal*.

<https://doi.org/10.2139/ssrn.1411462>

- Dennis, P., Mayhew, S., & Stivers, C. (2006). Stock Returns, Implied Volatility Innovations, and the Asymmetric Volatility Phenomenon. *The Journal of Financial and Quantitative Analysis*, 41(2), 381–406. JSTOR.
- Hibbert, A. M., Daigler, R. T., & Dupoyet, B. (2008). A behavioral explanation for the negative asymmetric return–volatility relation. *Journal of Banking & Finance*, 32(10), 2254–2266. <https://doi.org/10.1016/j.jbankfin.2007.12.046>
- Jiang, G. J., & Tian, Y. S. (2006). Extracting Model-Free Volatility from Option Prices: An Examination of the Vix Index. *SSRN Electronic Journal*.
<https://doi.org/10.2139/ssrn.880459>
- Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263. <https://doi.org/10.2307/1914185>
- Koenker, R. (2005). *Quantile regression*.
- Koenker, R., & Bassett, G. (1982). Robust Tests for Heteroscedasticity Based on Regression Quantiles. *Econometrica*, 50(1), 43–61. JSTOR.
<https://doi.org/10.2307/1912528>
- Koenker, R., & Hallock, K. F. (2001). Quantile Regression. *Journal of Economic Perspectives*, 15(4), 143–156. <https://doi.org/10.1257/jep.15.4.143>
- Rubinstein, M. (1994). Implied Binomial Trees. *The Journal of Finance*, 49(3), 771–818. <https://doi.org/10.1111/j.1540-6261.1994.tb00079.x>
- Sharpe, W. F. (1964). CAPITAL ASSET PRICES: A THEORY OF MARKET EQUILIBRIUM UNDER CONDITIONS OF RISK*. *The Journal of Finance*, 19(3), 425–442. <https://doi.org/10.1111/j.1540-6261.1964.tb02865.x>
- Simon, D. P. (2003a). The Nasdaq Volatility Index During and After the Bubble. *The Journal of Derivatives*, 11(2), 9–24. <https://doi.org/10.3905/jod.2003.319213>

- Simon, D. P. (2003b). The Nasdaq Volatility Index During and After the Bubble. *The Journal of Derivatives*, *11*(2), 9–24. <https://doi.org/10.3905/jod.2003.319213>
- Wu, G. (2001). The Determinants of Asymmetric Volatility. *The Review of Financial Studies*, *14*(3), 837–859. <https://doi.org/10.1093/rfs/14.3.837>
- Xu, X., & Taylor, S. J. (1995). Conditional volatility and the informational efficiency of the PHLX currency options market. *Journal of Banking & Finance*, *19*(5), 803–821. [https://doi.org/10.1016/0378-4266\(95\)00086-V](https://doi.org/10.1016/0378-4266(95)00086-V)

Appendix

```
1. from ast import If
2. import numpy as np
3. import pandas as pd
4. import datetime as dt
5.
6.
7. yields = pd.read_csv('12M RF.csv', sep=';', usecols=
    [11,12])
8. yields['OBS_VALUE'] = yields['OBS_VALUE']
9.
10. path = "obx_raw.csv"
11. df = pd.read_csv(
12.     path,
13.     sep="," ,
14.     header=None,
15.     parse_dates=[0],
16.     date_parser=lambda col: pd.to_datetime(col,
    format="%Y%m%d" ,),
17.     low_memory=False
18. )
19.
20. df.drop([1,3,4,6,7,10,11,12,13,14,15,16], axis=1,
    inplace=True)
21.
22. df.columns = ['Date', 'Option', 'Expiration', 'Bid',
    'Ask']
23.
24. df['Expiration'] =
    pd.to_datetime(df['Expiration'],format="%Y%m%d")
25. df['Days'] = df['Expiration'] - df['Date']
26.
27. df['Days'] = pd.to_numeric(df['Days'].dt.days,
    downcast='integer')
28.
29. result = df['Option'].str.split('(\d+)([A-Za-z]+)',
    expand=True)
30. result = result.drop([0,1,4,5,6], axis=1)
31. result = result.loc[:, [2,3]]
32. result.rename(columns={2:'Option type', 3:'Strike'},
    inplace=True)
33.
34. df = pd.concat([df,result],axis=1)
35.
36. df['Strike'] = pd.to_numeric(df['Strike'], errors
    ='coerce')
37. df.drop(['Option','Expiration'], axis=1, inplace=True)
38. df.dropna(subset = ['Strike'], inplace=True)
39.
40. c = ('A', 'B', 'C', 'D', 'E', 'F', 'G', 'H', 'I', 'J',
    'K', 'L')
```

```

41. p = ('M', 'N', 'O', 'P', 'Q', 'R', 'S', 'T', 'U', 'V',
      'W', 'X')
42. c1 = []
43. p1 = []
44.
45. c1.append(df[df['Option type'].isin(c)])
46. p1.append(df[df['Option type'].isin(p)])
47.
48. calls = pd.concat(c1)
49. puts = pd.concat(p1)
50. calls.drop(['Option type'], axis=1, inplace=True)
51. puts.drop(['Option type'], axis=1, inplace=True)
52.
53. calls.rename(columns={'Bid': 'Call_Bid', 'Ask':
      'Call_Ask'}, inplace=True)
54. puts.rename(columns={'Bid': 'Put_Bid', 'Ask':
      'Put_Ask'}, inplace=True)
55.
56. CP_df = calls.merge(puts, on=['Date', 'Strike',
      'Days'], how='outer')
57.
58. CP_df['Call_Premium'] = (CP_df['Call_Bid'] +
      CP_df['Call_Ask']) / 2
59. CP_df['Put_Premium'] = (CP_df['Put_Bid'] +
      CP_df['Put_Ask']) / 2
60. CP_df.loc[CP_df['Call_Bid'] < 0, 'Call_Premium'] = np.nan
61. CP_df.loc[CP_df['Put_Bid'] < 0, 'Put_Premium'] = np.nan
62. CP_df.dropna(subset = ['Call_Premium', 'Put_Premium'],
      inplace=True)
63.
64. CP_df.rename(columns={'Call_Premium': 'C', 'Put_Premium':
      'P'}, inplace=True)
65. CP_df.set_index(['Date', 'Days', 'Strike'], inplace=True)
66.
67. CP_df = CP_df.reset_index().set_index(['Date', 'Days',
      'Strike']).sort_index()
68.
69.
70. CP_df['CPdiff'] = (CP_df['C'] - CP_df['P']).abs()
71.
72. CP_df['min'] = CP_df['CPdiff'].groupby(level =
      ['Date', 'Days']).transform(lambda x: x == x.min())
73.
74.
75. df_all = CP_df.reset_index().set_index(['Date',
      'Days']).sort_index()
76.
77.
78. df = CP_df[CP_df['min'] == 1].reset_index()
79.
80. yields.rename(columns={'TIME_PERIOD': 'Date',
      'OBS_VALUE': 'Rate'}, inplace=True)
81. yields['Date'] = pd.to_datetime(yields['Date'])
82. df = pd.merge(df, yields.reset_index(), how = 'left')

```

```

83.
84. df.drop(['index'], axis=1, inplace=True)
85.
86.
87. df['Forward'] = df['CPdiff'] * np.exp(df['Rate'] *
      df['Days']/36500)
88. df['Forward'] += df['Strike']
89.
90. forward = df[['Date', 'Days', 'Forward']]
91.
92. df.set_index(['Date', 'Days'], inplace=True)
93.
94.
95. df_all = df_all.merge(forward, on=['Date', 'Days'], how
      = 'left')
96. df_all.set_index(['Date', 'Days'], inplace=True)
97. print(df_all)
98.
99. df_all['mid_strike'] = df_all[df_all['Strike'] <
      df_all['Forward']]['Strike'].groupby(level =
      ['Date', 'Days']).max()
100. df29 = df_all.reset_index()
101. mid_strike = df29[['Date', 'Days', 'mid_strike']]
102.
103.
104. df_all.loc[df_all['mid_strike']<0, 'mid_strike'] =
      np.nan
105. df_all.dropna(subset = ['mid_strike'], inplace=True)
106.
107.
108. df_puts = df_all.drop(['Call_Bid', 'Call_Ask', 'C'],
      axis=1)
109. df_calls = df_all.drop(['Put_Bid', 'Put_Ask', 'P'],
      axis=1)
110.
111. df_calls = df_calls.assign(zero_bid=lambda df:
      (df_calls['Call_Bid'] == 0).astype(int))
112. df_calls['zero_bid_accum'] = df_calls.groupby(level =
      ['Date', 'Days'])['zero_bid'].cumsum()
113.
114. df_puts = df_puts.groupby(level =
      ['Date', 'Days']).apply(lambda x:
      x.sort_values(['Strike'], ascending = False))
115. df_puts = df_puts.assign(zero_bid=lambda df:
      (df_puts['Put_Bid'] == 0).astype(int))
116. df_puts['zero_bid_accum'] = df_puts.groupby(level =
      ['Date', 'Days'])['zero_bid'].cumsum()
117.
118. df_puts = df_puts.reset_index().set_index(['Date',
      'Days']).sort_index()
119. df_calls = df_calls.reset_index().set_index(['Date',
      'Days']).sort_index()
120.

```



```

121.df_puts = df_puts[(df_puts['zero_bid_accum'] < 2) &
    (df_puts['Put_Bid'] > 0)]
122.df_calls = df_calls[(df_calls['zero_bid_accum'] < 2) &
    (df_calls['Call_Bid'] > 0)]
123.
124.
125.
126.condition1 = df_all['Strike'] < df_all['mid_strike']
127.condition2 = df_all['Strike'] > df_all['mid_strike']
128.
129.df_all['Premium'] = (df_all['P'] + df_all['C']) / 2
130.
131.df_all['Premium'].loc[condition1] =
    df_all['P'].loc[condition1]
132.df_all['Premium'].loc[condition2] =
    df_all['C'].loc[condition2]
133.
134.print(df_all)
135.
136.def compute_adjoning_strikes_diff(group):
137.     new = group.copy()
138.     if len(new) >= 3:
139.         new.iloc[1:-1] = np.array((group.iloc[2:] -
            group.iloc[:-2]) / 2)
140.         new.iloc[0] = group.iloc[1] - group.iloc[0]
141.         new.iloc[-1] = group.iloc[-1] - group.iloc[-2]
142.         return new
143.     else:
144.         new.iloc[:] = np.nan
145.         return new
146.
147.df_all['dK'] = df_all.groupby(level = ['Date',
    'Days'])['Strike'].apply(compute_adjoning_strikes_diff)
148.df_all = pd.merge(df_all.reset_index(),
    yields.reset_index(), how = 'left')
149.
150.df_all['sigma2'] = df_all['dK'] / df_all['Strike'] ** 2
151.df_all['sigma2'] *= df_all['Premium'] *
    np.exp(df_all['Rate'] * df_all['Days'] / 36500)
152.
153.print(df_all)
154.
155.sigma2 =
    df_all.groupby(['Date', 'Days'])[['sigma2']].sum() * 2
156.
157.sigma2 = sigma2.merge(mid_strike, on=['Date', 'Days'],
    how = 'left')
158.sigma2 = sigma2.merge(forward, on=['Date', 'Days'], how
    = 'left')
159.
160.sigma2.set_index(['Date', 'Days'], inplace=True)
161.
162.sigma2['sigma2'] -= (sigma2['Forward'] /
    sigma2['mid_strike'] - 1) ** 2

```

```

163. sigma2['sigma2'] /=
    sigma2.index.get_level_values(1).astype(float) / 365
164. sigma2 = sigma2[['sigma2']]
165.
166. sigma2 = sigma2.reset_index()
167.
168.
169. print(sigma2)
170.
171. def f(group):
172.     days = np.array(group['Days'])
173.     sigma2 = np.array(group['sigma2'])
174.
175.     if days.min() <= 30:
176.         T1 = days[days <= 30].max()
177.     else:
178.         T1 = days.min()
179.
180.     if len(days[days > T1]) > 0:
181.         T2 = days[days > T1].min()
182.     else:
183.         T2 = days.min()
184.
185.     sigma_T1 = sigma2[days == T1][0]
186.     sigma_T2 = sigma2[days == T2][0]
187.
188.     return pd.DataFrame([{'T1' : T1, 'T2' : T2,
        'sigma2_T1' : sigma_T1, 'sigma2_T2' : sigma_T2}])
189.
190. two_sigmas =
    sigma2.reset_index().groupby('Date').apply(f).groupby(lev
    el = 'Date').first()
191.
192. print(two_sigmas)
193.
194.
195. df = two_sigmas.copy()
196.
197.
198.
199. for t in ['T1', 'T2']:
200.     df['days_' + t] = df[t].astype(float) / 365
201.     df[t] = (df[t] - 1) * 1440. + 510 + 930
202.
203.
204. df['sigma2_T1'] = df['sigma2_T1'] * df['days_T1'] *
    (df['T2'] - 30. * 1440.)
205. df['sigma2_T2'] = df['sigma2_T2'] * df['days_T2'] *
    (30. * 1440. - df['T1'])
206. df['VIX'] = ((df['sigma2_T1'] + df['sigma2_T2']) /
    (df['T2'] - df['T1']) * 365. / 30.) ** .5 * 100
207.
208. VIX = df[['VIX']]
209.

```

```

210. print(VIX)
211.
212.
213.     import pandas as pd
214.     from scipy.stats import pearsonr
215.     from functools import reduce
216.     import numpy as np
217.     import matplotlib.pyplot as plt
218.     import seaborn as sns
219.
220.
221.     sp = pd.read_csv("SP500_return.csv", usecols =
        ['Exchange Date', 'Close', '%Chg'])
222.     sp.columns = ['Date', 'SP', 'ChgSP']
223.     sp["Date"] = pd.to_datetime(sp["Date"])
224.     sp = sp.loc[sp['Date'] <= '2020-05-22']
225.     sp.drop([0], axis = 0, inplace = True)
226.     sp['SP'] = sp['SP'].str.replace(" ", "")
227.     sp['SP'] = pd.to_numeric(sp['SP'])
228.
229.
230.
231.     obx = pd.read_csv("OBX_return.csv", usecols =
        ['Exchange Date', 'Close', '%Chg'])
232.     obx.columns = ['Date', 'OBX', 'ChgOBX']
233.     obx["Date"] = pd.to_datetime(obx["Date"])
234.     obx = obx.loc[obx['Date'] <= '2020-05-22']
235.     obx.drop([0,1], axis = 0, inplace = True)
236.
237.
238.     vix = pd.read_csv("vix.csv", usecols =
        ['Close', 'Date'])
239.     vix.columns = ['Date', 'VIX']
240.     vix = vix.loc[vix['Date'] <= '2020-05-22']
241.     vix["Date"] = pd.to_datetime(vix["Date"])
242.     #print(vix)
243.
244.     obx_vix = pd.read_csv('OBX_VIX_FINAL.csv', usecols
        = ['Date', 'OBX_VIX'])
245.     obx_vix["Date"] = pd.to_datetime(obx_vix["Date"])
246.     obx_vix.drop([0], axis = 0, inplace = True)
247.     #print(obx_vix)
248.
249.     data_frames = [obx_vix, vix, obx, sp]
250.     df = reduce(lambda left, right:
        pd.merge(left, right, on=['Date'],

```

```

251.
    how='outer'), data_frames)
252.
253.
254.     df.sort_values(by=['Date'], inplace = True)
255.     df.drop([0], axis = 0, inplace = True)
256.     df.dropna(subset = ['OBX_VIX', 'OBX', 'VIX', 'SP'],
    inplace=True)
257.     df.drop(['ChgOBX', 'ChgSP'], axis = 1 ,
    inplace=True)
258.
259.     df['OBX_VIXchg'] = np.log(df.OBX_VIX) -
    np.log(df.OBX_VIX.shift(1))
260.     df['VIXchg'] = np.log(df.VIX) -
    np.log(df.VIX.shift(1))
261.     df['OBXchg'] = np.log(df.OBX) -
    np.log(df.OBX.shift(1))
262.     df['SPchg'] = np.log(df.SP) -
    np.log(df.SP.shift(1))
263.     df.drop([1], axis = 0, inplace = True)
264.
265.     selected_col = df[['Date', 'OBX_VIXchg', 'VIXchg',
    'OBXchg', 'SPchg']]
266.     df_chg = selected_col.copy()
267.
268.
269.     fig,ax = plt.subplots()
270.     ax.plot(df['Date'], df['OBX_VIX'], color="red",
    linewidth=1, label = "OBX VIX")
271.     ax.plot(df['Date'], df['VIX'], color="green",
    linewidth=1, label = "VIX")
272.     ax.set_xlabel("Date",fontsize=14)
273.     ax.set_ylabel("Volatility Index",fontsize=14)
274.
275.     ax2=ax.twinx()
276.     ax2.plot(df['Date'], df['OBX'], linewidth=1,
    color="blue", label = "OBX")
277.
278.     ax2.set_ylabel("OBX",fontsize=14)
279.     lines_labels = [ax.get_legend_handles_labels() for
    ax in fig.axes]
280.     lines, labels = [sum(lol, []) for lol in
    zip(*lines_labels)]
281.
282.     fig.legend(lines, labels)
283.
284.     fig.set_size_inches(10, 7.5)

```

```

285.
286.     plt.show()
287.
288.
289.     #Correlation for the entire sample
290.     corr = pearsonr(df['VIX'], df['OBX_VIX'])
291.     print("Correlation for the entire sample VIX and
        OBX_VIX:", corr)
292.
293.     #Correlation for first sample (insert time)
294.
295.     sample1 = df.loc[df['Date'] <= '2010-12-31']
296.     corrs1 = pearsonr(sample1['VIX'],
        sample1['OBX_VIX'])
297.     print("Correlation for the first sample VIX and
        OBX_VIX:", corrs1)
298.
299.     #Correlation for second sample (insert time)
300.
301.     sample2 = df.loc[(df['Date'] >= '2011-01-01') &
        (df['Date'] <= '2015-12-31')]
302.     corrs2 = pearsonr(sample2['VIX'],
        sample2['OBX_VIX'])
303.     print("Correlation for the second sample VIX and
        OBX_VIX:", corrs2)
304.
305.     #Correlation for third sample (insert time)
306.
307.     sample3 = df.loc[df['Date'] >= '2016-01-01']
308.     corrs3 = pearsonr(sample3['VIX'],
        sample3['OBX_VIX'])
309.     print("Correlation for the third sample VIX and
        OBX_VIX:", corrs3)
310.
311.     #Correlation for US
312.         #Correlation for entire sample
313.     uscorr = pearsonr(df['SPchg'], df['VIXchg'])
314.     print("Correlation for the entire sample S&P and
        VIX:", uscorr)
315.         #First sample
316.     us1 = df.loc[df['Date'] <= '2010-12-31']
317.     uscorr1 = pearsonr(us1['SPchg'], us1['VIXchg'])
318.     print("Correlation for the first sample S&P and
        VIX:", uscorr1)
319.
320.         #Second sample

```

```

321.     us2 = df.loc[(df['Date'] >= '2011-01-01') &
                    (df['Date'] <= '2015-12-31')]
322.     uscorr2 = pearsonr(us2['SPchg'], us2['VIXchg'])
323.     print("Correlation for the second sample S&P and
            VIX:", uscorr2)
324.
325.         #Third sample
326.     us3 = df.loc[df['Date'] >= '2016-01-01']
327.     uscorr3 = pearsonr(us3['SPchg'], us3['VIXchg'])
328.     print("Correlation for the third sample S&P and
            VIX:", uscorr3)
329.
330.     #Correlation for Norway
331.         #Correlation for the entire sample
332.     nocorr = pearsonr(df['OBXchg'], df['OBX_VIXchg'])
333.     print("Correlation for the entire sample OBX and
            OBX_VIX:", nocorr)
334.
335.         #First sample
336.     no1 = df.loc[df['Date'] <= '2010-12-31']
337.     nocorr1 = pearsonr(no1['OBXchg'],
                        no1['OBX_VIXchg'])
338.     print("Correlation for the first sample OBX and
            OBX_VIX:", nocorr1)
339.
340.         #Second sample
341.     no2 = df.loc[(df['Date'] >= '2011-01-01') &
                    (df['Date'] <= '2015-12-31')]
342.     nocorr2 = pearsonr(no2['OBXchg'],
                        no2['OBX_VIXchg'])
343.     print("Correlation for the second sample OBX and
            OBX_VIX:", nocorr2)
344.
345.         #Third sample
346.     no3 = df.loc[df['Date'] >= '2016-01-01']
347.     nocorr3 = pearsonr(no3['OBXchg'],
                        no3['OBX_VIXchg'])
348.     print("Correlation for the third sample OBX and
            OBX_VIX:", nocorr3)
349.
350.     #Correlation between OBX and S&P 500
351.         #Entire sample
352.     incorr = pearsonr(df['OBXchg'], df['SPchg'])
353.     print("Correlation for the entire sample OBX and
            S&P:", incorr)
354.
355.         #First sample

```

```

356.     in1 = df.loc[df['Date'] <= '2010-12-31']
357.     incorr1 = pearsonr(in1['OBXchg'], in1['SPchg'])
358.     print("Correlation for the first sample OBX and
           S&P:", incorr1)
359.
360.         #Second sample
361.     in2 = df.loc[(df['Date'] >= '2011-01-01') &
           (df['Date'] <= '2015-12-31')]
362.     incorr2 = pearsonr(in2['OBXchg'], in2['SPchg'])
363.     print("Correlation for the second sample OBX and
           S&P:", incorr2)
364.
365.         #Third sample
366.     in3 = df.loc[df['Date'] >= '2016-01-01']
367.     incorr3 = pearsonr(in3['OBXchg'], in3['SPchg'])
368.     print("Correlation for the third sample OBX and
           S&P:", incorr3)
369.
370.     # Correlation betweenn OBX and VIX
371.         #Entire sample
372.     ovcorr = pearsonr(df['OBXchg'], df['VIXchg'])
373.     print("Correlation for the entire sample OBX and
           VIX:", ovcorr)
374.
375.         #First sample
376.     ov1 = df.loc[df['Date'] <= '2010-12-31']
377.     ovcorr1 = pearsonr(ov1['OBXchg'], ov1['VIXchg'])
378.     print("Correlation for the first sample OBX and
           VIX:", ovcorr1)
379.
380.         #Second sample
381.     ov2 = df.loc[(df['Date'] >= '2011-01-01') &
           (df['Date'] <= '2015-12-31')]
382.     ovcorr2 = pearsonr(ov2['OBXchg'], ov2['VIXchg'])
383.     print("Correlation for the second sample OBX and
           VIX:", ovcorr2)
384.
385.         #Third sample
386.     ov3 = df.loc[df['Date'] >= '2016-01-01']
387.     ovcorr3 = pearsonr(ov3['OBXchg'], ov3['VIXchg'])
388.     print("Correlation for the third sample OBX and
           VIX:", ovcorr3)
389.
390.     ""
391.     Calculating realized 30 days volatility
392.
393.     ""

```

```

394.
395.     window = 21
396.     trd_days = 252
397.     ann_factor = trd_days
398.
399.     df['rlz_var_sp'] = df['SPchg']**2
400.     df['rlz_var_sp'] =
        df['rlz_var_sp'].rolling(window).sum()
401.     df['rlz_vol_sp'] = np.sqrt(df['rlz_var_sp']) *
        np.sqrt(trd_days)
402.
403.     df['rlz_var_obx'] = df['OBXchg']**2
404.     df['rlz_var_obx'] =
        df['rlz_var_obx'].rolling(window).sum()
405.     df['rlz_vol_obx'] = np.sqrt(df['rlz_var_obx']) *
        np.sqrt(trd_days)
406.
407.     #Correlation realized volatility OBX and S&P
408.     df.dropna(subset = ['rlz_vol_sp', 'rlz_vol_obx'],
        inplace=True)
409.
410.     #Entire sample
411.     rlz_corr = pearsonr(df['rlz_vol_sp'],
        df['rlz_vol_obx'])
412.     print("Correlation entire sample realized
        volatility OBX and S&P:", rlz_corr)
413.
414.     #First sample
415.     rlz1 = df.loc[df['Date'] <= '2010-12-31']
416.     rlz_corr1 = pearsonr(rlz1['rlz_vol_sp'],
        rlz1['rlz_vol_obx'])
417.     print("Correlation first sample realized volatility
        OBX and S&P:", rlz_corr1)
418.
419.     #Second sample
420.     rlz2 = df.loc[(df['Date'] >= '2011-01-01') &
        (df['Date'] <= '2015-12-31')]
421.     rlz_corr2 = pearsonr(rlz2['rlz_vol_sp'],
        rlz2['rlz_vol_obx'])
422.     print("Correlation second sample realized
        volatility OBX and S&P:", rlz_corr2)
423.
424.     #Third sample
425.     rlz3 = df.loc[df['Date'] >= '2016-01-01']
426.     rlz_corr3 = pearsonr(rlz3['rlz_vol_sp'],
        rlz3['rlz_vol_obx'])

```



```

427.     print("Correlation third sample realized volatility
          OBX and S&P:", rlz_corr3)
428.
429.     #Correlation realized volatility and implied
          volatility
430.     df['rlz_vol_obx'] = df['rlz_vol_obx'] * 10
431.
432.     #Entire sample
433.     rlziv_corr = pearsonr(df['OBX_VIX'],
          df['rlz_vol_obx'])
434.     print("Correlation entire sample realized
          volatility and implied volatility OBX and OBX_VIX:",
          rlziv_corr)
435.
436.     #First sample
437.     rlziv1 = df.loc[df['Date'] <= '2010-12-31']
438.     (rlziv1['OBX_VIX'], rlziv1['rlz_vol_obx'])
439.     print("Correlation first sample realized volatility
          and implied volatility OBX and OBX_VIX:", rlziv_corr1)
440.
441.     #Second sample
442.     rlziv2 = df.loc[(df['Date'] >= '2011-01-01') &
          (df['Date'] <= '2015-12-31')]
443.     rlziv_corr2 = pearsonr(rlziv2['OBX_VIX'],
          rlziv2['rlz_vol_obx'])
444.     print("Correlation second sample realized
          volatility and implied volatility OBX and OBX_VIX:",
          rlziv_corr2)
445.
446.     #Third sample
447.     rlziv3 = df.loc[df['Date'] >= '2016-01-01']
448.     rlziv_corr3 = pearsonr(rlziv3['OBX_VIX'],
          rlziv3['rlz_vol_obx'])
449.     print("Correlation third sample realized volatility
          and implied volatility OBX and OBX_VIX:", rlziv_corr3)
450.
451.
452.     fig1,ax1 = plt.subplots()
453.
454.     ax1.plot(df['Date'], df['OBX_VIX'], color="red",
          linewidth=1, label = "IV OBX")
455.     ax1.plot(df['Date'], df['rlz_vol_obx'],
          color="green", linewidth=1, label = "RV OBX")
456.
457.     ax1.set_xlabel("Date",fontsize=14)
458.     ax1.set_ylabel("Volatility" ,fontsize=14)
459.

```

```

460.     lines_labels1 = [ax1.get_legend_handles_labels()
        for ax1 in fig1.axes]
461.     lines1, labels1 = [sum(lol, []) for lol in
        zip(*lines_labels1)]
462.
463.     fig1.legend(lines1, labels1)
464.
465.     fig1.set_size_inches(10, 7.5)
466.
467.     plt.show()
468.
469.     turnover = pd.read_excel("Contract turnover 1990-
        2020.xls.xlsx", usecols = [1,3])
470.     turnover = turnover.loc[23:37]
471.     turnover.columns = ['Date', 'Turnover']
472.     turnover['Turnover'] = turnover['Turnover']/1000000
473.     fig2, ax2 = plt.subplots()
474.
475.     ax2.bar(turnover['Date'], turnover['Turnover'])
476.     ax2.set_xlabel("Date", fontsize=14)
477.     ax2.set_ylabel("Turnover in million contracts"
        , fontsize=14)
478.
479.
480.     fig2.set_size_inches(10, 7.5)
481.
482.     plt.show()
483.
484. from itertools import groupby
485. import numpy as np
486. import pandas as pd
487. from matplotlib import pyplot as plt
488. from sklearn.linear_model import LinearRegression
489. from sklearn.model_selection import KFold
490. import statsmodels.api as sm
491. from statsmodels.stats.diagnostic import het_white
492. import statsmodels.formula.api as smf
493. from sklearn import linear_model
494. from scipy.stats import kurtosis
495. from scipy.stats import skew
496.
497.
498. obxvix = pd.read_csv('OBX_VIX_FINAL.csv', usecols =
        ['Date', 'OBX_VIX'])
499. obxvix.columns = ['Date', 'Vix']
500.
501.
502. obxvix['Date'] = pd.to_datetime(obxvix['Date'])
503.
504.

```

```

505.dataobx = pd.read_csv('OBX_return.csv', usecols =
    ['Exchange Date', 'Close'])
506.
507.dataobx.rename(columns={'Exchange Date': 'Date',
    'Close': 'Prices'}, inplace=True)
508.dataobx['Date'] = pd.to_datetime(dataobx['Date'])
509.
510.dataobx['OBX_returns'] = np.log(dataobx.Prices) -
    np.log(dataobx.Prices.shift(1))
511.dataobx = dataobx.dropna()
512.
513.
514.df1 = dataobx.merge(obxvix, on=['Date'], how='inner')
515.
516.
517.fig,ax = plt.subplots()
518.ax.plot(df1.Date, df1.Prices, color='red')
519.ax.set_xlabel('year', fontsize=10)
520.ax.set_ylabel('Prices', fontsize=10)
521.
522.ax2=ax.twinx()
523.ax2.plot(df1.Date, df1['Vix'], color='blue')
524.ax2.set_ylabel('Vix', fontsize=10)
525.plt.show()
526.
527.df1['SMA5_OBX'] = df1['Prices'].rolling(5).mean()
528.df1['DevMA5'] = (df1['SMA5_OBX'] - df1['Prices']) /
    df1['Prices']
529.
530.df1['Vix_first_diff']= df1['Vix'].diff()
531.df1['Vix_returns']= np.log(df1.Vix) -
    np.log(df1.Vix.shift(1))
532.
533.df1 = df1.dropna()
534.
535.
536.#x_1 = df1.OBX_returns
537.#y_1 = df1.Vix_returns
538.
539.#print(y)
540.
541.#x = sm.add_constant(x_1) #Adds the constant
542.#modell=sm.OLS(y,x).fit() #Makes the regression model
    with the defined y and x
543.#modell.summary()
544.#print(modell.summary())
545.#plt.scatter(x_1, y)
546.#yhat1 = 0.0023 -0.3489 * x_1
547.#fig1 = plt.plot(x_1,yhat1, c="red", linewidth=2)
548.#plt.show()
549.
550.#White's test
551.#white_test = het_white(modell.resid,
    modell.model.exog)

```

```

552.
553. #define labels to use for output of White's test
554. #labels = ['Test Statistic', 'Test Statistic p-value',
             'F-Statistic', 'F-Test p-value']
555.
556. #print results of White's test
557. #print(dict(zip(labels, white_test)))
558.
559.
560. df2 = pd.DataFrame()
561.
562.
563. df2 = df2.append(df1.Vix_returns)
564. df2 = df2.append(df1.OBX_returns)
565. df2 = df2.T
566. df2['OBXpos'] =
    df1.OBX_returns.where(df1.OBX_returns>=0,0)
567. df2['OBXneg'] =
    df1.OBX_returns.where(df1.OBX_returns<0,0)
568. df2['SMA5pos'] = df1.DevMA5.where(df1.DevMA5>=0,0)
569. df2['SMA5neg'] = df1.DevMA5.where(df1.DevMA5<0,0)
570. df2['VIX_first_diff'] = df1.Vix_first_diff
571. df2['Vix_return_lagged'] = df1['Vix_returns'].shift(1)
572.
573.
574. df2 = df2.dropna()
575.
576.
577. '''
578. print('Mean Vix:')
579. print(df1['Vix'].mean())
580. print('Mean Vix returns:')
581. print(df1['Vix_returns'].mean()*100)
582.
583. print('Std Vix:')
584. print(df1['Vix'].std())
585. print('Std Vix returns:')
586. print(df1['Vix_returns'].std()*100)
587.
588.
589. print('Vix min:')
590. print(df1['Vix'].min())
591. print('Vix returns min:')
592. print(df1['Vix_returns'].min()*100)
593.
594.
595. print('Vix max:')
596. print(df1['Vix'].max())
597. print('Vix_returns max:')
598. print(df1['Vix_returns'].max()*100)
599.
600. print('Vix skew:')
601. print(df1['Vix'].skew())
602. print('Vix return skew:')

```

```

603.print(df1['Vix_returns'].skew())
604.
605.print('Vix kurt:')
606.print(kurtosis(df1['Vix'], fisher = True, bias =
        False))
607.print('Vix returns kurt:')
608.print(kurtosis(df1['Vix_returns'], fisher = True, bias
        = False))
609. '''
610.
611.#Multivariate regression (Model 1)
612.X_mv1 = df2[['Vix_return_lagged']]
613.y_mv1 = df2['VIX_first_diff']
614.
615.xx1 = sm.add_constant(X_mv1)
616.model_mult1 = sm.OLS(y_mv1,xx1).fit()
617.model_mult1.summary()
618.print(model_mult1.summary())
619.
620.#Multivariate regression (Model 2)
621.X_mv2 = df2[['Vix_return_lagged', 'OBXpos']]
622.y_mv2 = df2['VIX_first_diff']
623.
624.xx2 = sm.add_constant(X_mv2)
625.model_mult2 = sm.OLS(y_mv2,xx2).fit()
626.model_mult2.summary()
627.print(model_mult2.summary())
628.
629.#Multivariate regression (Model 3)
630.X_mv3 = df2[['Vix_return_lagged', 'OBXneg']]
631.y_mv3 = df2['VIX_first_diff']
632.
633.xx3 = sm.add_constant(X_mv3)
634.model_mult3 = sm.OLS(y_mv3,xx3).fit()
635.model_mult3.summary()
636.print(model_mult3.summary())
637.
638.#Multivariate regression (Model 4)
639.X_mv4 = df2[['Vix_return_lagged', 'OBXpos', 'SMA5pos']]
640.y_mv4 = df2['VIX_first_diff']
641.
642.xx4 = sm.add_constant(X_mv4)
643.model_mult4 = sm.OLS(y_mv4,xx4).fit()
644.model_mult4.summary()
645.print(model_mult4.summary())
646.
647.#Multivariate regression (Model 5)
648.X_mv5 = df2[['Vix_return_lagged', 'OBXneg', 'SMA5neg']]
649.y_mv5 = df2['VIX_first_diff']
650.
651.xx5 = sm.add_constant(X_mv5)
652.model_mult5 = sm.OLS(y_mv5,xx5).fit()
653.model_mult5.summary()
654.print(model_mult5.summary())

```

```

655.
656. #Multivariate regression (Model 6)
657. X_mv5 = df2[['Vix_return_lagged', 'OBXpos', 'OBXneg',
        'SMA5pos', 'SMA5neg']]
658. y_mv5 = df2['VIX_first_diff']
659.
660. xx5 = sm.add_constant(X_mv5)
661. model_mult5 = sm.OLS(y_mv5,xx5).fit()
662. model_mult5.summary()
663. print(model_mult5.summary())
664.
665.
666. #Quantile regression
667. mod = smf.quantreg("Vix_returns ~ OBX_returns", df2)
668. res = mod.fit(q=0.5)
669. #print(res.summary())
670.
671. quantiles = np.arange(0.1, 1, 0.1)
672.
673. def fit_model(q):
674.     res =mod.fit(q=q)
675.     return [q, res.params['Intercept'],
        res.params['OBX_returns']] + res.conf_int().loc[
676.         'OBX_returns'
677.     ].tolist()
678.
679. models = [fit_model(x) for x in quantiles]
680. models = pd.DataFrame(models, columns =['q', 'a', 'b',
        'lb', 'ub'])
681.
682. ols = smf.ols('Vix_returns ~ OBX_returns', df2).fit()
683. ols_ci = ols.conf_int().loc['OBX_returns'].tolist()
684. ols = dict(
685.     a=ols.params['Intercept'],
        b=ols.params['OBX_returns'], lb=ols_ci[0], ub=ols_ci[1]
686.)
687.
688. print(models)
689. print(ols)
690.
691. #Plot quantile regression and OLS
692. x = np.arange(df2.OBX_returns.min(),
        df2.OBX_returns.max(), 0.01)
693. get_y = lambda a, b: a + b * x
694.
695. fig, ax3 = plt.subplots(figsize=(8,6))
696.
697. for i in range(models.shape[0]):
698.     y = get_y(models.a[i], models.b[i])
699.     ax3.plot(x, y, linestyle='dotted', color='grey')
700.
701. y = get_y(ols['a'], ols['b'])
702.
703. ax3.plot(x, y, color='red', label='OLS')

```

```

704. ax3.scatter(df2.OBX_returns, df2.Vix_returns, alpha
              =0.5)
705. ax3.set_xlim((-0.15,0.15))
706. ax3.set_ylim((-0.5,0.5))
707. legend = ax3.legend()
708. ax3.set_xlabel('OBX_returns', fontsize=16)
709. ax3.set_ylabel('Vix_returns', fontsize=16)
710. plt.show()
711.
712. #second plot
713. n = models.shape[0]
714. p1 = plt.plot(models.q, models.b, color='black',
                 label='Quantile Reg. ')
715. p2 = plt.plot(models.q, models.ub, linestyle='dotted',
                 color='black')
716. p3 = plt.plot(models.q, models.lb, linestyle='dotted',
                 color='black')
717. p4 = plt.plot(models.q, [ols['b']] * n, color='red',
                 label='OLS')
718. p5 = plt.plot(models.q, [ols['lb']] * n,
                 linestyle='dotted', color='red')
719. p6 = plt.plot(models.q, [ols['ub']] * n, linestyle=
                 'dotted', color='red')
720. plt.ylabel(r"$\beta_{OBX\_returns}$")
721. plt.xlabel('Quantiles of the conditional Vix ret
                 distribution')
722. plt.legend()
723. plt.show()
724.
725.
726. #Multiple Quantile Regression
727.
728. mod2 = smf.quantreg("Vix_returns ~ 1 + OBXpos +
                       OBXneg", df2)
729. res2 = mod2.fit(q=0.5)
730.
731.
732. from ast import If
733. from heapq import merge
734. import numpy as np
735. import pandas as pd
736. import datetime as dt
737. from itertools import groupby
738. from matplotlib import pyplot as plt
739. from sklearn.linear_model import LinearRegression
740. from sklearn.model_selection import KFold
741. import statsmodels.api as sm
742. from statsmodels.stats.diagnostic import het_white
743. import statsmodels.formula.api as smf
744. from sklearn import linear_model
745. from scipy.stats import kurtosis
746. from scipy.stats import skew

```

```

747.     from functools import reduce
748.     import seaborn as sns
749.     from sklearn import datasets
750.     from sklearn.linear_model import LinearRegression
751.
752.
753.     obxvix = pd.read_csv('OBX_VIX_FINAL.csv', usecols =
        ['Date', 'OBX_VIX']) #2007 til 2020
754.     obxvix.columns = ['Date', 'Vix']
755.     obxvix.rename(columns={'Vix': 'obxvix_prices'},
        inplace=True) #obxvix_prices
756.     obxvix['Date'] = pd.to_datetime(obxvix['Date'])
757.
758.     obxvix['obxvix_returns'] =
        np.log(obxvix.obxvix_prices) -
        np.log(obxvix.obxvix_prices.shift(1))
759.     obxvix = obxvix.dropna()
760.
761.     dataobx = pd.read_csv('OBX_return.csv')
762.
763.     dataobx.drop(['%Chg', 'Unnamed: 3', 'Unnamed:
        4', 'Unnamed: 5', 'Unnamed: 6', 'Unnamed: 7', 'Unnamed:
        8', 'Unnamed: 9', 'Unnamed: 10', 'Unnamed: 11', 'Unnamed:
        12', 'Unnamed: 13', 'Unnamed: 14', 'Unnamed: 15'], axis=1,
        inplace=True)
764.     dataobx.rename(columns={'Exchange Date': 'Date',
        'Close': 'obx_prices'}, inplace=True)
765.     dataobx['Date'] = pd.to_datetime(dataobx['Date'])
766.
767.     dataobx['obx_returns'] = np.log(dataobx.obx_prices)
        - np.log(dataobx.obx_prices.shift(1))
768.     dataobx = dataobx.dropna()
769.
770.     sp500 = pd.read_csv('SP500_return.csv')
771.
772.     sp500.drop(['%Chg', 'Unnamed: 3', 'Unnamed:
        4', 'Unnamed: 5', 'Unnamed: 6', 'Unnamed: 7', 'Unnamed:
        8', 'Unnamed: 9', 'Unnamed: 10', 'Unnamed: 11', 'Unnamed:
        12'], axis=1, inplace=True)
773.     sp500.rename(columns={'Exchange Date': 'Date',
        'Close': 'sp500_prices'}, inplace=True)
774.     sp500['Date'] = pd.to_datetime(sp500['Date'])
775.
776.     sp500['SP500_returns'] = np.log(sp500.sp500_prices)
        - np.log(sp500.sp500_prices.shift(1))
777.     sp500 = sp500.dropna()
778.

```



```

779.     cbeovix= pd.read_csv('CBEOVIX.csv')
780.
781.     cbeovix.drop(['Unnamed: 2','Unnamed: 3','Unnamed:
      4','Unnamed: 5','Unnamed: 6','Unnamed: 7','Unnamed:
      8','Unnamed: 9','Unnamed: 10','Unnamed: 11','Unnamed:
      12'], axis=1, inplace=True)
782.     cbeovix.rename(columns={'Exchange Date': 'Date',
      'Close': 'cbeovix_prices'}, inplace=True)
783.     cbeovix['Date'] = pd.to_datetime(cbeovix['Date'])
784.
785.     cbeovix['cbeovix_returns'] =
      np.log(cbeovix.cbeovix_prices) -
      np.log(cbeovix.cbeovix_prices.shift(1))
786.     cbeovix = cbeovix.dropna()
787.
788.     '''
789.     #VIX DESCRIPTIVE
790.     print('Mean Vix:')
791.     print(cbeovix['cbeovix_prices'].mean())
792.     print('Mean Vix returns:')
793.     print(cbeovix['cbeovix_returns'].mean()*100)
794.
795.     print('Std Vix:')
796.     print(cbeovix['cbeovix_prices'].std())
797.     print('Std Vix returns:')
798.     print(cbeovix['cbeovix_returns'].std()*100)
799.
800.
801.     print('Vix min:')
802.     print(cbeovix['cbeovix_prices'].min())
803.     print('Vix returns min:')
804.     print(cbeovix['cbeovix_returns'].min()*100)
805.
806.
807.     print('Vix max:')
808.     print(cbeovix['cbeovix_prices'].max())
809.     print('Vix_returns max:')
810.     print(cbeovix['cbeovix_returns'].max()*100)
811.
812.     print('Vix skew:')
813.     print(cbeovix['cbeovix_prices'].skew())
814.     print('Vix return skew:')
815.     print(cbeovix['cbeovix_returns'].skew())
816.
817.     print('Vix kurt:')
818.     print(kurtosis(cbeovix['cbeovix_prices'], fisher =
      True, bias = False))

```

```

819.     print('Vix returns kurt:')
820.     print(kurtosis(cbeovix['cbeovix_returns'], fisher =
      True, bias = False))
821.     '''
822.
823.     data_frames = [dataobx, sp500, obxvix, cbeovix]
824.
825.     df = reduce(lambda left,right:
      pd.merge(left,right,on=['Date'],
826.             how='inner'), data_frames)
827.
828.     obxvixdata = dataobx.merge(obxvix, on=['Date'],
      how='inner')
829.
830.     obxvixdata['obxvix_first_diff'] =
      obxvixdata['obxvix_prices'].diff()
831.
832.     obxvixdata['obxvix_lagged'] =
      obxvixdata['obxvix_prices'].shift(1)
833.
834.     obxvixdata['VIXpos'] =
      obxvixdata.obxvix_returns.where(obxvixdata.obxvix_returns
      >=0,0).shift(1)
835.     obxvixdata['VIXneg'] =
      obxvixdata.obxvix_returns.where(obxvixdata.obxvix_returns
      <0,0).shift(1)
836.
837.     obxvixdata = obxvixdata.dropna()
838.
839.     #Split data into quantiles based on obxvix_prices
840.     q = obxvixdata.quantile([0.0, 0.1, 0.2, 0.3, 0.4,
      0.5, 0.6, 0.7, 0.8, 0.9, 1])
841.
842.     quantilerule = 'obxvix_returns'
843.
844.     q1 =
      obxvixdata[((obxvixdata[quantilerule]>=q[quantilerule][0.
      0]) & (obxvixdata[quantilerule]<q[quantilerule][0.1]))]
845.     q2 =
      obxvixdata[((obxvixdata[quantilerule]>=q[quantilerule][0.
      1]) & (obxvixdata[quantilerule]<q[quantilerule][0.2]))]
846.     q3 =
      obxvixdata[((obxvixdata[quantilerule]>=q[quantilerule][0.
      2]) & (obxvixdata[quantilerule]<q[quantilerule][0.3]))]

```

```

847.     q4 =
        obxvixdata[((obxvixdata[quantilerule]>=q[quantilerule][0.
            3]) & (obxvixdata[quantilerule]<q[quantilerule][0.4]))]
848.     q5 =
        obxvixdata[((obxvixdata[quantilerule]>=q[quantilerule][0.
            4]) & (obxvixdata[quantilerule]<q[quantilerule][0.5]))]
849.     q6 =
        obxvixdata[((obxvixdata[quantilerule]>=q[quantilerule][0.
            5]) & (obxvixdata[quantilerule]<q[quantilerule][0.6]))]
850.     q7 =
        obxvixdata[((obxvixdata[quantilerule]>=q[quantilerule][0.
            6]) & (obxvixdata[quantilerule]<q[quantilerule][0.7]))]
851.     q8 =
        obxvixdata[((obxvixdata[quantilerule]>=q[quantilerule][0.
            7]) & (obxvixdata[quantilerule]<q[quantilerule][0.8]))]
852.     q9 =
        obxvixdata[((obxvixdata[quantilerule]>=q[quantilerule][0.
            8]) & (obxvixdata[quantilerule]<q[quantilerule][0.9]))]
853.     q10 =
        obxvixdata[((obxvixdata[quantilerule]>=q[quantilerule][0.
            9]) & (obxvixdata[quantilerule]<q[quantilerule][1]))]
854.
855.
856.     #Calculate positive and negative OBX returns for
        each quantile
857.
858.     print('obx max:')
859.     print(q10['obxvix_returns'].max())
860.
861.     print(q10['obxvix_returns'].min())
862.     #Regression
863.     X = q10[['VIXpos', 'VIXneg']]
864.     y = q10['obx_returns']
865.
866.     xx = sm.add_constant(X)
867.     model_mult = sm.OLS(y,xx).fit()
868.     model_mult.summary()
869.     print(model_mult.summary())

```