

Abnormal Stock Market Returns around VIX Volatility Peaks

Is there evidence for abnormal return around volatility peaks?

Anette Dalen Brokke
Madeleine Østeby

Supervisor
Valeriy Zakamulin

This master's thesis is carried out as a part of the education at the University of Agder and is therefore approved as a part of this education. However, this does not imply that the University answers for the methods that are used or the conclusions that are drawn.

University of Agder, 2016
School of Business and Law
Department of Economics and Finance

Abstract

It has been observed that the volatility index, VIX, often ends in a spike peak. The intention when introducing VIX was to measure the volatility related to S&P indices. Evidence in the literature argues that VIX is an estimate of investors' sentiment and fear, by need for insurance of portfolio in terms of options. In the thesis we outlined three studies related to VIX and S&P 500. We perform a replica study of the already observed asymmetrical relation between VIX and S&P 500, where we include symmetrical models to discuss the results. Furthermore, two studies take the spike peak in VIX into account. The first determines the peaks of VIX in terms of rising and falling phases and contains statistical summary of the phases. The second study determines the abnormal return in S&P 500 around the peaks in VIX, with use of a window including ten days before peak and ten days after peak. The studies are motivated by the relationship between the VIX and S&P 500. There is limited literature about VIX in terms of phases of rising and falling volatility. The motivation for these studies is to find a clear structure of the relationship around the spike peak. To understand investors' sentiment, theory from the literature about behavioral finance is included. The research goals for the studies are to find clear patterns of how return acts in relation to volatility. We expect evidence of negative relationship between VIX and S&P 500, where S&P 500 increases more slowly than it drops. In terms of the abnormal returns we expect the return to be significantly negative before the peak and significantly positive in the days after the peak. The results from the replica study give evidence of a negative relationship, although we did not reach all the expected results as in Whaley (2000) and Whaley (2009). Results for the study of peaks show that duration of a falling phase is longer than duration of a rising phase, and the opposite for amplitude in VIX. Results for the abnormal study show a clear negative trend of abnormal return before the peak, cumulative abnormal return of -4.106 percent for the eleven first days of the event study. In addition, the abnormal return for the ten days after the peak gave 3.12 percent.

Acknowledgement

The thesis is the final part of our master degree program in finance at University of Agder, Kristiansand. The subject for our thesis is abnormal stock market return around VIX volatility peaks. We want to see if we can get expected results regarding the relationship between the change in VIX and the change in S&P 500. Further, we want to test for abnormal returns around peaks in VIX.

We will use the opportunity to thank our supervisor Professor Valeriy Zakamulin for motivating us during the whole process while he was writing an article about the same topic (Zakamulin, 2016). He has been of great help and guidance during the process, especially with the starting process regarding the statistical program R-project. Where he gave us the R files named “Identifying-Algorithms” and “Data”, these files are of limited changes and have been used as a basis for all other analysis in the thesis. We are also grateful to Rune Nilsson for constructive comments on the paper.

Table of Contents

1	Introduction.....	1
2	Literature Review	3
3	Behavioral Finance	6
4	Data	9
4.1	The VIX Index	9
4.2	Standard & Poor's 500 Index.....	10
4.3	Summary Statistics of S&P 500 Returns.....	11
5	The Relationship Between VIX and S&P 500	12
5.1	Methodology.....	12
5.1.1	Classical Linear Regression Model (CLRM).....	12
5.1.2	Ordinary Least Square (OLS)	12
5.2	Empirical Results	13
5.2.1	Symmetrical Regression	13
5.2.2	Asymmetric Regression	15
6	Study of Peaks in VIX	18
6.1	Methodology.....	18
6.1.1	Algorithm by Bry and Boschan	18
6.1.2	Algorithm by Lunde and Timmermann	19
6.1.3	Phase duration and amplitude	19
6.1.4	One sample t-test.....	20
6.1.5	Normality Test	21
6.1.6	Autocorrelation	22
6.1.7	Ljung-Box test.....	22
6.2	Empirical Results	23
6.2.1	Algorithm by Bry and Boschan	23
6.2.2	Algorithm by Lunde and Timmermann	25
7	Event Study of Abnormal Returns.....	26
7.1	Methodology.....	26
7.1.1	The Welch Two Sample t-test.....	26
7.1.2	Two Sample Wilcoxon Test.....	27
7.1.3	The Kolmogorov-Smirnov Two Sample Test.....	28
7.1.4	Event Study with Dummies	29

7.2	Empirical Results	31
7.2.1	Results of Mean Abnormal Return from Algorithm by Bry and Boschan	31
7.2.2	Event Study of Abnormal Returns from Algorithm by Bry and Boschan	32
7.2.3	Robustness Test from Algorithm by Bry and Boschan.....	35
7.2.4	Results of Mean Abnormal Return from Algorithm by Lunde and Timmermann	37
7.2.5	Event Study of Abnormal Returns from Algorithm by Lunde and Timmermann.....	38
7.2.6	Robustness Test from Algorithm by Lunde and Timmermann	40
8	Discussion	42
8.1	Future Research	45
9	Conclusion	46
10	References.....	47
11	Appendix.....	50
11.1	R-file	50
11.1.1	Filename: Download-Data.....	50
11.1.2	Filename: Identifying-Algorithms	50
11.1.3	Filename: Data	62
11.1.4	Filename: Algorithm-BB	65
11.1.5	Filename: Window.Table.Dummy.BB	73
11.1.6	Filename: Algorithm-LT	77
11.1.7	Filename: Window.Table.Dummy.LT.....	84
11.2	Reflection Note.....	89

List of tables

Table 4-1 Summary statistic for daily log-returns.	11
Table 5-1 Results of regression (5.2).....	14
Table 5-2 Report the model summary of regression (5.2).....	14
Table 5-3 Report the results of regression (5.3).	15
Table 5-4 Report the model summary of regression (5.3).	15
Table 5-5 Report the results of regression (5.4).	16
Table 5-6 Report the model summary of regression (5.4).	16
Table 5-7 Report the results of regression (5.5).	17
Table 5-8 Report the model summary of regression (5.5).	17
Table 6-1 Summary statistics of periods of rising and falling VIX using algorithm by BB. ..	24
Table 6-2 Summary statistics of periods of rising and falling VIX using algorithm by LT....	25
Table 7-1 Mean abnormal return around volatility peaks, algorithm by BB.....	31
Table 7-2 Mean abnormal return over the event window, algorithm by BB.....	33
Table 7-3 Mean abnormal return around volatility peaks, algorithm by LT.	37
Table 7-4 Mean abnormal return over the event window, algorithm by LT.	38

List of figures

Figure 4-1 S&P 500 and VIX.	9
Figure 7-1 Mean daily abnormal return over the event window, algorithm by BB.....	34
Figure 7-2 Mean cumulative abnormal return over the event window, algorithm by BB.....	34
Figure 7-3 Mean abnormal return for over the event window for full period, first half and second half of the sample, algorithm by BB.....	35
Figure 7-4 Cumulative abnormal return for over the event window for full period, first half and second half of the sample, algorithm by BB.	35
Figure 7-5 Mean abnormal return for over the event window for all peaks, above and below median amplitude of peaks, algorithm by BB.	36
Figure 7-6 Cumulative abnormal return for over the event window for all peaks, above and below median amplitude of peaks, algorithm by BB.....	37
Figure 7-7 Mean daily abnormal return over the event window, algorithm by LT.	39
Figure 7-8 Mean cumulative return over the event window, algorithm by LT.	39
Figure 7-9 Mean abnormal return for over the event window for full period, first half and second half of the sample, algorithm by LT.	40
Figure 7-10 Cumulative abnormal return for over the event window for full period, first half and second half of the sample, algorithm by LT.	40
Figure 7-11 Mean abnormal return for over the event window for all peaks, above and below median amplitude of peaks, algorithm by LT.....	41
Figure 7-12 Cumulative abnormal return for over the event window for all peaks, above and below median amplitude of peaks, algorithm by LT.	41

1 Introduction

Practitioners have observed that the stock market return is negative before the volatility peak, and positive after the peak. A spike peak in the stock market volatility is said to be a typical signal to buy stocks. The Chicago Board Options Exchange has reported the VIX index since 1993. The VIX index is a measure of implied volatility which is related to the S&P 500 index. The dynamics of VIX is characterized by many periodic peaks, and where a spike in VIX is related to a drop in the stock market price, S&P 500 index (Whaley, 2000). In the first article of VIX presented by Whaley (2000) the VIX index was called the “investor fear gauge”, this is because of the contemporaneous relationship between the VIX index and the S&P 500 index. Where its argued that VIX is an estimate of investors sentiment and fear. The thesis is motivated by the relationship between volatility and return in S&P 500. Where three studies lead up the how abnormal return in S&P500 act around spike peaks in volatility, measured by the VIX index. Included in the thesis are also theory about behavioral finance, to describe the results from the empirical studies by investors sentiment. The research goals are to find clear patterns in how abnormal return act in relation to volatility.

When studying the relationship between the VIX index and the S&P 500 index we replicated the studies done by Whaley (2000) and Whaley (2009), where we use asymmetric regressions as well as we included symmetric regressions. We looked at the symmetric and asymmetric relationship when changes S&P 500 predict the changes in VIX, and then we turned the equation around and looked at the relationship between the change in VIX predict the changes in S&P 500. Even though some of our results were unexpected, the relationship between the VIX index and the S&P 500 index is stronger when the market price decreases (VIX increases) compared to when the market price increases (VIX decreases). This relationship is asymmetric, and can be an explanation why the VIX index can be related to investors sentiment.

In the second study we used the algorithm developed by Bry and Boschan (1971) and Lunde and Timmermann (2003) to define turning points in bull and bear markets in the stock market, and determine peaks in the VIX index. We divided phases into rising and falling volatility. This topic is relatively new when using the volatility divided into rising and falling phases. The results showed that falling phases contains of longer periods than rising phases, while rising phases tends to increase more than the falling phases decrease. The third and the last study also

include peaks in VIX, we used the peaks we found in the second study. In this study we first test for mean abnormal return before and after a peak using 5, 10, 15, and 20 days. Second, we used the event study methodology with dummy variables to find if there is evidence of abnormal returns around peaks in VIX. It is not much existing literature on this topic, while regarding the event study methodology we used the event study presented by MacKinlay (1997) and Binder (1998). The study was done by using the full sample from January 1, 1990 to December 31, 2015, then we divided the full sample into two periods, first half (1990.01.01-2002.12.31) of the sample and second half (2003.01.01-2015.12.31) of the sample. Further we did a robustness test where we divided into above median amplitude and below median amplitude. The first part of this study showed that the mean abnormal return was negative before the peak, and positive after the peak. We found strong and robust evidence of abnormal returns around peaks in VIX.

In the next chapter we will present a literature review, then theory about behavioral finance and the data we have used, including a summary statistic. Further we will divide our methodology and empirical results into three sections where we in chapter 5 start with the study of the relationship between the VIX index and the S&P 500 index, chapter 6 the study of peaks in VIX, and chapter 7 the event study of abnormal returns around peaks in VIX. In chapter 8 we will discuss our findings from chapter 5-7 before we in chapter 9 concludes the thesis.

2 Literature Review

There are a few papers explaining how the VIX index is constructed and its statistical behavior. Some of the articles which describe the VIX index is Whaley (2000), Whaley (2009), and Fernandes, Medeiros, and Scharth (2014). Already in the very first article presented by Whaley (2000) the VIX index was called the “investor fear gauge”. He showed that the relationship between the change in stock market index and the changes in the VIX index is asymmetric. The regression showed that the stock market index is expected to increase by 0.775% if the VIX index does not change, due to premium for risky investments, the stock price is expected to increase over time. He included a new variable to make the regression asymmetric, which contains only positive changes in VIX. Further, the results state that if the VIX index decreases by 100 basis points, the stock market index will increase by 0.469%, and if the VIX index increases by 100 basis points, the S&P 100 index will decrease by -0.707%. The interdependence between the stock market reacts stronger to an increase in VIX than to a decrease in VIX. Thus, this asymmetry states that VIX is a measure of the investors fear rather than the expected future stock market volatility. Whaley (2009) use the relation between the VIX and the stock market, where he includes a new variable in the regression which contains only of returns when the market is going down to estimate the VIX. The intercept in his regression was insignificant which means if the stock market index does not change during the day, the change in the VIX index will be negligible. The relationship between the VIX index and the S&P 500 index is inverse and asymmetric. His results show if the S&P 500 index increases by 100 basis points, the VIX index will decrease by -2.99%, and if the S&P 500 index decreases by 100 basis points, the VIX index will increase by 4.493%. Fernandes, Medeiros, and Scharth (2014) and Giot (2005) research shows that the relationship between the VIX index and the S&P 500 index returns is strongly negative.

Whaley (2000) observed that the VIX index contains of many periodic peaks, and the high levels of volatility in the VIX index, the S&P 500 index will result in a drop in the stock market price. Some researchers noted that a spike peak in volatility index can be a buy signal for stocks because a peak in VIX is followed by a market rally. Results from Giot (2005) support the hypothesis about extremely high levels of volatility will be an attractive signal to buy stocks for traders which is in a long-term position. His study was based on indices, and not portfolios. Banerjee, Doran, and Peterson (2006) confirmed the results supported by Giot (2005). The VIX levels is related to future returns, while innovations are not related to future returns.

Some articles investigate whether the VIX index is a measure of investors sentiment. Copeland and Copeland (1999) states that the VIX index is a leading indicator for the daily stock market returns, and the market timing matters. His results show that the returns are different for indices associated with different size, values and growth, following the high levels of the VIX index. Brown and Cliff (2004) used two measures for their survey, individual investors and investors intelligence. Their results show that investors sentiment has little or no power for near-term future stock market returns. The strongest relationship they found was between institutional sentiment and large stocks, and for individual investors the sentiment is not limited. Baker and Wurgler (2006) also did a study based on investors sentiment. They wanted to challenge the classical finance theory which claim that investors sentiment has no effect on the cross-section when it comes to the stock market prices and returns. In the study they used simple theoretical arguments, historical speculative episodes, and a set of empirical results. Their results showed that investors sentiment has significant cross-sectional effects. Banerjee, Doran, and Peterson (2007) used models in behavioral finance to see if there is a relationship between the VIX levels and the returns. Their results suggest that it was not any difference whether the market was based on directional movements or volatility levels.

Several studies use bull and bear markets to explain phases of rising and falling prices in the market, where bull represents the rising period and bear the falling period in the market. Bry and Boschan (BB) (1971) developed an algorithm to define turning points in the business cycle. They also define a set of rules for determine the turning points, first the criterion for defining the location of a peak or trough, and second the duration between these points. The algorithm by BB has been adapted by other researcher. Cashin, McDermott, and Scott (2002) used the algorithm in the commodity price cycle. Pagan and Sossounov (2003) also applied a variant of the algorithm to find bull and bear markets using monthly stock market prices. When using smoothed data, important information could be lost because of the possible large movements in the equity markets, therefore they decided not to use smoothed data. Their results showed that the bull markets generally last longer than bear markets, and give higher returns. The bull markets result in an increase of 20% or more in the stock market price, while a minority of the bear markets result in a decrease of 20% or less. In the study done by Gonzalez, Powell, Shi, and Wilson (2003) they used the procedure of the BB algorithm where they used a time series of monthly stock market data. Their findings show that mean return shifts are associated with bull and bear market periods. Another algorithm to identify turning points in a time-series was presented by Lunde and Timmermann (LT) (2004). When selecting phases to determine bull

and bear markets they look at the changes in the stock market prices and how much the prices had change since the previous local peak or trough. The change in the stock market price had to be big enough to identify a new peak or trough. We explain the details in the algorithm by BB and the algorithm by LT later in this thesis.

3 Behavioral Finance

Related to finance theory are theory about behavioral finance. Usually in finance theory assumptions that the investor is rational are made, but in reality that is not always the case. When prices drop on the stock exchange, investors are afraid to lose and want to get out before it is too late, so they sell instead of buying. To understand behavioral finance; the market prices are driven by supply and demand by investors and how investors react to news in the market.

Investors are rational when they behave logically, they buy stocks when the market price decreases and sell stocks when the market price increases. In a rational world, trading should not be used too much. In the world today the volume of trading is too high related to rational models (Barberis & Thaler, 2003). The investors are not as rational as traditionally finance theory indicates because humans act on emotions (Baker & Nofsinger, 2010). Behavioral finance is a relatively new concept where investors seeks to use their emotions and cognitive psychological theory to make economic and financial conclusions. Behavioral finance is of interest in this thesis because it can be useful when explaining how the market is responding, and why we get the results we got.

The behavioral finance theory states that there exist several irrationalities which seems to have an impact on the investors decisions (Bodie, Kane, & Marcus, 2013). These irrationalities can be split in two categories, information processing and behavioral biases. The first category is mainly how investors can misunderstand certain information, and then the probability distributions about future returns will be misleading. The second category is about the investors having the correct probability distribution about future returns but make decisions which is not optimal or is inconsistent, which results in risk-return trade-offs (Bodie, Kane, & Marcus, 2013). The irrationalities will affect the efficiency of the market.

Fama and Malkiel (1970) came up with three different forms of markets being efficient, depending on the information available. First we have the weak-form efficient which includes all information about past prices. Second, the semi-strong efficient which includes all the public information in the market, and all the historical prices. Third, the strong-form efficient which includes all public and private information in the market. The semi-strong efficient is in relevance because we are using the event study methodology to test for this efficiency (Baker & Nofsinger, 2010). The efficient market hypothesis is an important statement before making

any conclusions. The efficient market hypothesis is where the aggregate supply curve and the aggregate demand curve crosses, which is where the price on a given security is equal to the expected value of all future cash flow at that given security. This can be violated in different forms. One violation is on investors own reaction on news events. Where one event can seem efficient in the short-run but when you study it in a long-run perspective it will not be as efficient. A reason is that investors react differently when they get “good news” and when they get “bad news”. The “good news” will drive the cumulative abnormal return up while “bad news” will drive the cumulative abnormal return down. This means that the market is changing also after the earnings announcement (Baker & Nofsinger, 2010).

Researchers have argued that investors being irrational, their biases will unlikely be systematic (Baker & Nofsinger, 2010). Each investor has different biases, and these biases will eventually cancel each other out. An unbiased rational investor should be able to take advantage of investors being systematically biased because of the biases, and this will drive out the investors not being rational out of the market. Investors mental functioning can be a struggle between two well-known principles, the pleasure principle and the reality principle. Regardless of how painful a decision can be, investors must have the ability to sense reality as it actually is, and not the way other people want it to be (Baker & Nofsinger, 2010).

Investors behavior may have different effects when they buy stocks and when they sell stocks (Barberis & Thaler, 2003). Several studies point out that investors are more reluctant to sell stocks when they know they are selling it for less than the price they were purchasing. This effect is called the disposition effect and was first explained by Shefrin and Statman (1985). Furthermore, Odean (1998) and Odean (1999) explains how investors typically acts when they must decide when to sell a stock and when to buy a stock. The selling decision says that investors are more willing to sell stocks which have increased in price rather than stocks which has decreased, relative to the price purchased. With this in mind the study shows on average that the stocks the investors sells are more valuable at that time than the stocks they are holding at that time, the investor are not that rational. The buying decision is affected more by an attention effect (Odean, 1999). According to the selling decision investors will gain some profit, this is not always the case when it comes to the buying decision. When investors buy stocks they are not always going through the list of all listed stocks but they are choosing a stock based on interest and their attention (Odean, 1999).

Shiller (2003) states that the market efficiency will be affected by the investors sentiment. This could be explained by two reasons; the first reason is that the anomalies discovered was often made by under reaction by investors as well as an overreaction. Second, the methodology will often be improved and time passes so the anomalies often disappear. He also mentioned there is some weakness about all theoretical models, as every model has properties of the ideal world but it is important to also see it from the actual world as it is today. Barberis and Thaler (2003) conclude that most of the findings about their study of financial behavior are narrow. The investors beliefs, or preferences, or their limits to arbitrage will have an effect on the theoretical models; not all of them together but each of them will still have an impact if they are standing alone. His conclusion state that in the future more and more theorist will take investors behavior to consideration in the models.

4 Data

The datasets used in this thesis are the CBOE Volatility Index and the Standard & Poor's 500 (S&P 500) index form 1st of January 1990 until 31st of December 2015.

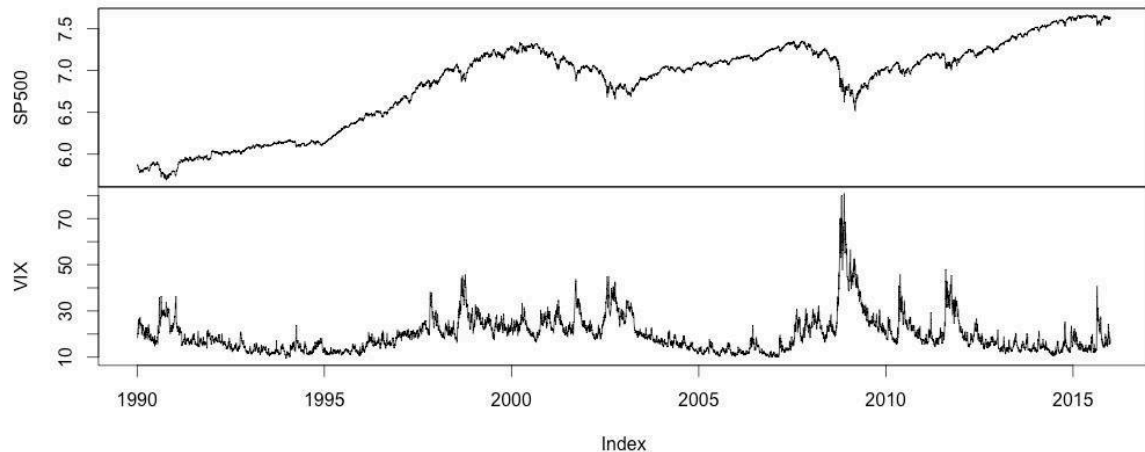


Figure 4-1 S&P 500 and VIX.

The data of our study. The top panel plots the log of the stock market index. The bottom panel plots the VIX volatility index. All data are in daily frequency from January 1, 1990 to December 31, 2015

4.1 The VIX Index

The CBOE Volatility Index, VIX, is a measure of implied volatility related to the S&P 500 Index. VIX was introduced on the Chicago Board Option Exchange in 1993, the dating of the index is calculated back to the beginning of 1990 so investors could compare the VIX Index with the previous method of calculating volatility, the VOX Index. Therefore, the data frame in this thesis cannot go further back than 1990. The VIX Index is a measure of the 30-day future volatility of the S&P 500 Index. The VIX Index is calculated of option prices which is forward looking rather than stock prices which is backward looking. The option prices are a reflection of the market's expected future volatility. In the calculation of VIX the call and put options with more than 23 and less than 37 days to expiration are used (Chicago Board Options Exchange). In early stage of VIX, calculating in-the-money and at-the-money options of the S&P 100 index (OEX) was used, as S&P 100 index had a higher turnover then the S&P 500. In September 2003 the VIX index started to use options related to S&P 500 (SPX) as well as out-of-the-money options. The introduction of out-of-the-money options was included in the calculation due to its consistence of important information about the investors needs for insurance. The

investors demand for out-of-the-money put options shows a demand for insurance of investments, as the demand increases so does the VIX (Whaley, 2009). The formula for generalized computation of VIX is as follows,

$$VIX = \sigma * 100,$$

where σ is computed as,

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

where T represent time to expiration, F is the forward index level from index option p , K_0 is the first strike below the forward index level F , and K_1 is the strike price of out of the money option (Chicago Board Options Exchange).

The calculation of the VIX index is similar to the calculation of internal rate of return. The same principal matter for calculation of VIX, the VIX value implies the current price of S&P 500 index options. And represents the expected volatility over the next 30 days (Whaley, 2009). The VIX index is a forward looking index, it measures the volatility the investors are expecting to see in the future. A backward looking index would measure the already realized volatility. In this thesis the VIX index ^VIX quote from Yahoo Finance is used (Yahoo!, 2016).

4.2 Standard & Poor's 500 Index

The S&P 500 index are a combination of the 500 largest publicly traded companies in the United States. The S&P 500 Index is an indicator of the total market even though the stocks included are in the large cap segment. The index is under the Index Committee, a team of S&P Dow Jones Indices economists and index analysts. They are responsible for maintenance and to keep S&P 500 the leading proxy of the market (S&P Dow Jones Indices , 2016). In this thesis the S&P 500 ^GSPC quote from Yahoo Finance is used (Yahoo!, 2016).

4.3 Summary Statistics of S&P 500 Returns

Statistics	1990-2015	1990-2002	2003-2015
Minimum	-9.4695	-7.1127	-9.4695
Mean, μ	0.0265	0.0273	0.0248
p-value of $H_0: \mu = 0$	0.0589	0.1374	0.2433
Standard deviation	1.1359	1.0512	1.2139
Maximum	10.9572	5.5744	10.9572
Skewness	-0.2399	-0.1166	-0.3207
Kurtosis	8.6505	3.8156	11.0342
Shapiro-Wilk normality test, p-value	0.0000	0.0000	0.0000
Autocorrelation, ρ_1	-0.0540	0.0100	-0.1020
p-value of $H_0: \rho_1 = 0$	0.0014	0.5664	0.0000

Table 4-1 Summary statistic for daily log-returns.

Minimum, mean, standard deviation, and maximum are reported in percentages.

The summary statistics is presented in table 3.1 for the full sample and two sub periods, one for the first half and one for the second half of the sample. The returns in the data sample are computed on the basis of the following formula,

$$r_t = \ln\left(\frac{p_t}{p_{t-1}}\right),$$

where r_t is the return at time t , and p_t is the price of the S&P 500 at time t .

The daily mean return is quite stable across the periods, while the standard deviation of the second period was higher than in the first period. Skewness is negative across all periods while the kurtosis is positive. We can see from the table that the first half of the period is closer a normal distribution than the full period and the second period. The Shapiro-Wilk test is rejecting the null hypothesis about the normal distribution of the sample in all of the periods. This is also the case when we test the null hypothesis $H_0: \mu = 0$ using a one sample t-test, which states that in the full sample as well as in the two sub periods, the daily mean return is not statistically significantly different from zero when using a 5% significance level. Table 3.1 also reports the value of the first-order autocorrelation for daily mean returns as well as the p-value for the Ljung-Box test, testing $H_0: \rho_1 = 0$. The full sample, as well as the second period, is not autocorrelated using a 5% significance level, while the second period is autocorrelated while failing to reject a 5% significance level.

5 The Relationship Between VIX and S&P 500

In the study of the relationship between VIX and S&P 500 we replicate the study of Whaley (2000) and Whaley (2009). His study is pioneering within the topic, where he looks at the asymmetric relationship. In this thesis we have also included the symmetric relationship.

5.1 Methodology

5.1.1 Classical Linear Regression Model (CLRM)

The classical linear regression model can be expressed as,

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t, \quad (5.1)$$

where y_t is the dependent variable at time t (time index), β_0 is the constant, β_1 is the slope of the independent variable x_t , and ε_t is the disturbance term at time t . Model (4.1) is based on five assumptions that must hold. First, the mean of the disturbance term is equal to zero ($E(\varepsilon_t) = 0$). Second, the assumption of homoscedasticity, the variance of the disturbance term is constant, finite, independent and identically distributed ($var(\varepsilon_t) = \sigma^2 < \infty$). Third, the covariance between the disturbance term is equal to zero ($cov(\varepsilon_i, \varepsilon_j) = 0$ for $i \neq j$), the covariance is linearly independent and uncorrelated. Fourth, the covariance between the error and the x variable is equal to zero, thus there is no relationship. Fifth, it is required that the disturbance term is normally distributed ($\varepsilon_t \sim N(0, \sigma^2)$) (Brooks, 2008).

5.1.2 Ordinary Least Square (OLS)

In this thesis the ordinary least squared (OLS) method is used to estimate the parameters of the model. OLS is a method of linear regression where it wants to find a best fit of a line to the data sample. When assumption 1-4 above holds the OLS is efficient (Brooks, 2008). In this thesis it is performed two symmetric and two asymmetric OLS regression models to explain the relation between the log-returns of the S&P 500 index and the changes in the VIX volatility index, as in Whaley (2000) and Whaley (2009). The difference of Whaley's research and the research outlined in this thesis are the time-frame and that the study in the thesis uses log returns. The regressions are expressed as follows,

$$r_t = \beta_0 + \beta_1 \Delta VIX_t + \varepsilon_t, \quad (5.2)$$

$$\Delta VIX_t = \beta_0 + \beta_1 r_t + \varepsilon_t, \quad (5.3)$$

$$r_t = \beta_0 + \beta_1 \Delta VIX_t + \beta_2 \Delta VIX_t^+ + \varepsilon_t, \quad (5.4)$$

$$\Delta VIX_t = \beta_0 + \beta_1 r_t + \beta_2 r_t^- + \varepsilon_t, \quad (5.5)$$

where r_t represents the log-return of the S&P 500 Index at time t (time index) and ΔVIX_t denotes the change in VIX at time t , β_0 is the intercept, β_1 is the slope of the first independent variable, for the asymmetric regressions β_2 is the slope of the second independent variable and ε_t denotes the disturbance term. In the first asymmetric regression ΔVIX_t^+ denotes the positive values of change in VIX and are equal to ΔVIX_t if $\Delta VIX_t > 0$, and zero otherwise (Whaley, 2000). The second asymmetric regression includes r_t^- which is equal to r_t if $r_t < 0$, and zero otherwise (Whaley, 2009). The null hypothesis and the alternative hypothesis will be expressed as,

$$H_0: \beta_i = 0,$$

$$H_1: \beta_i \neq 0.$$

The null hypothesis states that there is not a causality between the log-returns of the S&P 500 and the change in the VIX index. We reject the null hypothesis when the p-value is less than the chosen level of significance. If we reject the null hypothesis, we support the alternative hypothesis which says that there is a causality between the log-returns of the S&P 500 and the change in the VIX index

5.2 Empirical Results

The empirical results include the regression for both symmetrical and asymmetrical relation.

5.2.1 Symmetrical Regression

To find if there is a causality between the changes in the VIX index and the level of log-returns in the S&P 500 index we use the OLS regression (5.2), where the log-returns is the dependent variable (r_t) and the changes in VIX is the independent variable (ΔVIX_t). Then we use the OLS regression (5.3), where the changes in VIX is the dependent variable and the log-returns is the

independent variable. The two regressions are symmetric and regression (5.2) assumes the ΔVIX has equal impact of S&P 500 on rising and falling periods in the ΔVIX index. Regression (5.3) assumes that negative and positive log-returns has equal impact on the ΔVIX index.

The results from the regression in R:

$$r_t = \beta_0 + \beta_1 \Delta VIX_t + \varepsilon_t$$

Coefficient:				
	Estimate	Std. error	t-value	Pr(> t)
β_0	5.158e-04	9.901e-05	5.21	1.95e-07 ***
β_1	-1.247e-01	1.532e-03	-81.38	< 2e-16 ***

Significant code: '***'0.001

Table 5-1 Results of regression (5.2)

<i>Residual standard error</i>	<i>0.00801 on 6550 degrees of freedom</i>
<i>Multiple R-squared</i>	<i>0.5028</i>
<i>Adjusted R-squared</i>	<i>0.5027</i>
<i>F-statistic</i>	<i>6623 on 1 and 6550 DF</i>
<i>p-value</i>	<i>< 2.2e-16</i>

Table 5-2 Report the model summary of regression (5.2)

$$r_t = 0.0005158 - 0.1247$$

The symmetric regression shows that there is a causality between the changes in VIX and the log-returns of the S&P 500. First, both the constant (β_0) and the independent variable (β_1) is statistically significant different from zero at a 99% confidence interval because their p-values is lower than 0.01. The intercept is statistically different from zero, which means if the ΔVIX is not changing the S&P 500 would still grow. The results are not surprising; the stocks are expected to grow through time in order to compensate investors for risky investments. The relationship between the daily change in VIX and the daily change in S&P 500 are as followed, if ΔVIX increase by one percent the S&P 500 index decrease by -0.1247 percent. Second, the F statistics indicates that one or more variable can explain the daily return at a 99% confidence level, also shown by the low p-value. The R-squared is 0.5028 and explains 50.28% of the variance. In both the t-test and F-test it is evidence to reject of the null hypothesis and support the alternative hypothesis.

The results from the regression in R:

$$\Delta VIX_t = \beta_0 + \beta_1 r_t + \varepsilon_t$$

	Coefficient:			
	Estimate	Std. error	t-value	Pr(> t)
β_0	0.0030789	0.0005628	5.47	4.66e-08 ***
β_1	-4.0317169	0.0495399	-81.38	< 2e-16 ***

Significant code: '***'0.001.

Table 5-3 Report the results of regression (5.3).

<i>Residual standard error</i>	<i>0.045551 on 6550 degrees of freedom</i>
<i>Multiple R-squared</i>	<i>0.5028</i>
<i>Adjusted R-squared</i>	<i>0.5027</i>
<i>F-statistic</i>	<i>6623 on 1 and 6550 DF</i>
<i>p-value</i>	<i>< 2.2e-16</i>

Table 5-4 Report the model summary of regression (5.3).

$$\Delta VIX_t = 0.0030789 - 4.0317169$$

Both estimates are statistically significantly different from zero at a 99% confidence interval because the p-value is less than 0.01. We reject the null hypothesis and states that there is a causality between the daily change in VIX and the daily log-returns of S&P 500. If S&P 500 increase by one percent, the ΔVIX index will decrease by -4.0317169 percent. We found it unexpected that the intercept is statistically significant since ΔVIX is not expected to increase over time.

5.2.2 Asymmetric Regression

To test for the significant results with the asymmetric regression we start by using the regression (5.4) where a new variable, ΔVIX_t^+ , are included in the regression. As in the to the study done by Whaley (2000) where he regresses the r_t impact on ΔVIX . In the regression he includes a new variable which contains of all values which has a positive change in VIX. This is replica of his study, to predict r_t in the terms of a rising change in VIX should have more impact on r_t . It is expected that the variable with positive change in VIX has a significant negative impact on the log-return.

The results from the regression in R:

$$r_t = \beta_0 + \beta_1 \Delta VIX_t + \beta_2 \Delta VIX_t^+ + \varepsilon_t$$

	Coefficient:			
	Estimate	Std. error	t-value	Pr(> t)
β_0	-9.719e-05	1.436e-04	-0.677	0.499
β_1	-1.414e-01	3.229e-03	-43.797	< 2e-16 ***
β_2	2.705e-02	4.599e-03	5.882	4.26e-09 ***

Significant code: '***' 0.001.

Table 5-5 Report the results of regression (5.4).

<i>Residual standard error</i>	<i>0.00799 on 6549 degrees of freedom</i>
<i>Multiple R-squared</i>	<i>0.5054</i>
<i>Adjusted R-squared</i>	<i>0.5052</i>
<i>F-statistic</i>	<i>3346 on 2 and 6549 DF</i>
<i>p-value</i>	<i>< 2.2e-16</i>

Table 5-6 Report the model summary of regression (5.4).

$$r_t = -0.00009719 - 0.1414\Delta VIX_t + 0.002705\Delta VIX_t^+$$

The result from the new variable is not as expected. The ΔVIX_t^+ variable are statistically significant at a 99% confidence interval because its p-values is less than 0.01, but with a positive impact on the return. It is to notice that the estimate for the ΔVIX variable has decreased. The betas for independent variables are statistically significant on a 99% confidence level and R-squared has a slightly higher value the symmetrical regression (5.2). As mentioned before, the stock return is expected to increase over time. In this model the intercept is negative but not significant in the model. With our time frame and log-return we did not achieve similar results as Whaley (2009) did in his study of the impact ΔVIX and ΔVIX^+ have on the return in S&P 500.

Next we wanted to test for the significant results with the asymmetric regression using the regression (5.5) where the new variable, r_t^- , are included in the regression. This variable contains of all daily negative log-returns. It is expected that r_t^- have a negative impact on ΔVIX . Since r_t^- contains only negative values, ΔVIX is expected to increase more with the variable included.

The results from regression in R:

$$\Delta VIX_t = \beta_0 + \beta_1 r_t + \beta_2 r_t^- + \varepsilon_t$$

	Coefficient:			
	Estimate	Std. error	t-value	Pr(> t)
β_0	-0.0034725	0.0007558	-4.594	4.42e-06 ***
β_1	-3.1500683	0.0845087	-37.275	< 2e-16 ***
β_2	-1.7023846	0.1330371	-12.796	< 2e-16 ***

Significant code: '***'0.001.

Table 5-7 Report the results of regression (5.5).

<i>Residual standard error</i>	<i>0.007988 on 6549 degrees of freedom</i>
<i>Multiple R-squared</i>	<i>0.5056</i>
<i>Adjusted R-squared</i>	<i>0.5054</i>
<i>F-statistic</i>	<i>3349 on 2 and 6549 DF</i>
<i>p-value</i>	<i>< 2.2e-16</i>

Table 5-8 Report the model summary of regression (5.5).

$$\Delta VIX_t = -0.0034725 + -3.1500683r_t + -1.7023846r_t^- + \varepsilon_t$$

The results from the regression is quite similar to the results done by Whaley (2009). The exception is that the intercept is significant, meaning the ΔVIX index is to fall if there is no daily change in return. The explanation can be that log-returns are used in this thesis. On the other hand, the regression clearly states that the daily return and the negative daily return have an impact on the change in VIX. With 99% confidence level the variables states that ΔVIX will increase more if the return is negative. The R-squared is higher than the R-squared of the corresponding symmetric regression. If the log-returns of S&P 500 would decrease by 100 basis points, the ΔVIX index will rise by

$$\Delta VIX_t = -0.0035 - 3.1501(-0.1) + 1.7024(-0.1) = 0.4888\%$$

while, if the log-returns of S&P 500 would increase by 100 basis points, the ΔVIX index will fall by

$$\Delta VIX_t = -0.0035 - 3.1501(0.1) + 1.7024(0) = -0.3185\%$$

This study give us a good indications of the causality of the relationship using the symmetric regressions, while using the asymmetric regression which has a higher R^2 we get a better explanation of the relationship. Testing the relationship between the change in S&P 500 and the change in VIX gives us insight in how the two indexes react to each other, this is good knowledge when going further to the next studies in this thesis.

6 Study of Peaks in VIX

The purpose of the study of peaks in VIX is to study the relationship with the log-return in S&P 500. Other studies have studied the relationship between VIX and return in S&P 500, but not in terms of rising and falling volatility phases. It has been observed that a rising volatility-phase often ends in a spike peak (Whaley, 2000).

6.1 Methodology

To identify the turning points in the VIX Index, the phases of rising and falling volatility have to be identified. The shift from rising to falling volatility-phase is where the peaks are located, and the shift from falling to rising volatility-phases is where the troughs are located. To identify the phases needed for the turning points two different algorithms from financial literature of identifying bull and bear markets are used. In financial markets there is a unison agreement that a rising market is a bull state and a falling market is a bear state. One definition of a bull market, if the price increase or decrease by a more than a given percent since the latest local trough or peak. Another if the duration of an increasing market is longer than a given time frame, with more than a given movement in the index. The algorithms for this propose are the Algorithm by Bry and Boschan (BB) and the Algorithm by Lunde and Timmermann (LT). The time frame of a bull or bear markets can be over a period of many years. In this thesis the time frame used for rising and falling phases in the VIX Index is in days, that means the thesis look at the short term effect of a peak in the VIX Index have on the log-return.

6.1.1 Algorithm by Bry and Boschan

The first method to identify rising and falling volatility-phases is the Algorithm by BB (1971). The algorithm they developed seeks to isolate pattern in a business cycle by a set of rules, and are relevant for all time-series. The method smoothes the data, then set criterions for detecting peaks and troughs, criterions for the duration of a phases and cycles and criterion for movement in the index value. According to the paper by Pagan and Sossounov (2003) they successfully implemented the method to work for stock price data. By remove smoothing the data from the original algorithm, interesting and important data of a spike peak will still be in touch. In analysis of economic data extreme values from a maxima or a minima, and what happens around that point is the most interesting information. Since the thesis have a short term view, these extreme values are important to the analysis. In view of the short term effect of economic

data and that we are interested in the spike peak observed in the VIX index, the reshaped version is a good fit to detecting turning points of volatility. To determine turning points the algorithm sets a window of length l_{window} on either side of the peak or trough to find out if the VIX value is higher or lower than the VIX values around. A margin size on either side is used to eliminate local minimum or maximum unless the change is larger than margin size. Two criteria for length is used, one for the phase length and one for the full cycle. In addition, the threshold, θ , of relative change in VIX invalidates the length criteria if the change exceeds θ . Numbers used in the identification of phases are as follows, l_{window} equals 4, margin size equals 20 percent, duration of a cycle equals 22 days, duration of a phase 10 days and threshold equals 30 percent.

6.1.2 Algorithm by Lunde and Timmermann

The second method to identify rising and falling volatility-phases is the algorithm by LT (2004). According to their definition of the algorithm, the market switches from bull to bear when the price decreases beyond a given percentage, threshold, since the latest local peak within the bull market. The threshold is set to 30 percent change in VIX for determine bull. The opposite for a switch from bear to bull market, the switch takes place when the price increase beyond a given percentage, threshold, since the latest local trough within the bear state. The threshold for determine bear state is set to 20 percent.

6.1.3 Phase duration and amplitude

After identifying the peaks and troughs there are two main factors with information of the peak, duration and amplitude. Both factors are calculated from the peaks and troughs after the use of algorithm by BB or algorithm by LT. The duration displays information of how long the rising or falling phases are. A long rising duration in the VIX index implies that no new and sudden positive information are released to the market, which affects the volatility to significantly decrease. Or opposite for a decreasing duration. The formula for rising duration are as follows,

$$D = t_{\text{peak}} - t_{\text{trough}}, \quad (6.1)$$

where D denotes the duration in business days, t_{peak} denotes the observation of a peak and t_{trough} denotes the corresponding observation of a trough. On the other hand, a long duration

says nothing about the increase or decrease in the VIX Index value. To find the height of the peak is relative to the corresponding trough, amplitude is used. The formula for rising amplitude,

$$A_{Rising} = \frac{VIX_{peak} - VIX_{trough}}{VIX_{trough}}, \quad (6.2)$$

where A_{Rising} denotes the amplitude of the rising volatility-phase, VIX_{peak} denotes the VIX value of a peak and VIX_{trough} denotes the VIX value of the corresponding trough. To find the falling amplitude of how low a trough is relative to the peak the following formula is used,

$$A_{Falling} = \frac{VIX_{trough} - VIX_{peak}}{VIX_{peak}}, \quad (6.3)$$

where $A_{Falling}$ denotes the amplitude of falling volatility-phases.

The amplitude for both rising and falling says something about the relative height of the peak. A hypothesis can be drawn that the larger the gap from peak to trough or trough to peak the more impact have the switch on the S&P 500 Index.

6.1.4 One sample t-test

Sometimes we only want to perform a one sample t-test, because we only have one-sample data. The one sample t-test assumes the sample to be normally distributed ($N(\mu, \sigma^2)$) with mean μ and standard deviation σ^2 . In this test we want to test if the sample mean μ is different from a specific value, μ_0 , (Dalgaard, 2008). The hypothesis will be expressed as follows (Taeger & Kuhnt, 2014),

$$H_0: \mu = \mu_0,$$

$$H_1: \mu \neq \mu_0,$$

where the null hypothesis states that the sample mean is equal to the specific value, and the data are normally distributed. The alternative hypothesis states that the data is not

normally distributed because the sample mean and the specific value is different. The test statistic (Taeger & Kuhnt, 2014) is as follows,

$$t = \frac{\bar{\mu} - \mu_0}{s}, \quad (6.4)$$

where

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (\mu_i - \bar{\mu})^2}, \quad (6.5)$$

and t follows a student t-distribution with degrees of freedom of,

$$df = n - 1. \quad (6.6)$$

To find our p-value we can use a standard table of t-values where we use our test statistic value and our degrees of freedom. We can compare the value we find in the table with our chosen level of significance. Or if the p-value is less than the significance level, we reject the null hypothesis.

6.1.5 Normality Test

There are different ways to test whether our data follows a normal distribution. A simple test for the normality assumption is made by Jarque and Bera (1987) where they use the OLS residuals skewness and kurtosis. A normal distribution is not skewed (*skewness* = 0) and has a kurtosis of 3. Skewness is a measure of the asymmetry about its mean value and kurtosis measure the thickness of the tails of its distribution (Brooks, 2008).

There are other formal tests we can perform to check for normality. We used the Shapiro-Wilk normality test where we can test random samples to see whether they follow a normal distribution. The test statistic can be expressed as (Shapiro & Wilk, 1965),

$$W = \frac{(\sum_{i=1}^n \alpha_i x_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (6.7)$$

where α_i represents the constant generated from its mean and covariance of order statistics, x_i is the order statistics, \bar{x} is the sample mean, and n represents the sample size. The hypothesis can be expressed as,

$$H_0: x \sim \mathcal{N}(0, \sigma^2),$$

$$H_1: x \not\sim \mathcal{N}(0, \sigma^2),$$

where the null hypothesis states that the sample is normally distributed, whereas the alternative hypothesis say that the sample is not normally distributed. The null hypothesis can be rejected if the p-value is less than the chosen level of significance.

6.1.6 Autocorrelation

One assumption when using the OLS method is that the disturbance term is not correlated with other disturbance. This means that we want the disturbance term in our model to be uncorrelated (Brooks, 2008).

To test for autocorrelation, we test if there exist a relationship between the current value and the previous value (lagged value). While using an autocorrelation test, the sample disturbance will not be observed so the test is conducted using the residuals. Two type of autocorrelation, positive autocorrelation which is when the residuals at the previous period ($t - 1$) is positive, the residual at the current value (t) is also being positive, and the other way around. And negative autocorrelation where the residual at the previous period ($t - 1$) is positive while the residual at the current value (t) is negative, and opposite (Brooks, 2008). The equation can be expressed as,

$$y_t = \beta_0 + \beta_1 y_{t-1} + \varepsilon_t, \quad (6.8)$$

where y_{t-1} represents the previous lagged value.

6.1.7 Ljung-Box test

Another way to test for autocorrelation among the residuals is to use the modified test created by Ljung and Box (1978). The null hypothesis states that all of the lagged values m (in our case we deal with only one lagged value) of the correlation coefficients r_k are equal to zero. The test statistic can be expressed as (Ljung & Box, 1978),

$$\tilde{Q}(\hat{r}) = n(n + 2) \sum_{k=1}^m \frac{\hat{r}_k^2}{n-k}, \quad (6.9)$$

where n is the sample size. The hypothesis of the test is,

$$H_0: r_k = 0,$$

$$H_1: r_k \neq 0.$$

We can reject the null hypothesis if the p-value from this test is lower than the chosen level of significance. When we reject the null hypothesis it means that the lagged values of the correlation coefficient are not equal to zero, and it exist of autocorrelation in the residuals of the model.

6.2 Empirical Results

The peaks of the VIX Index are the top values of the index found by the sort of either the algorithm by BB or the algorithm by LT. The peak is there due to a shift from rising to falling volatility-phases.

6.2.1 Algorithm by Bry and Boschan

The statistics related to the algorithm by BB shows that there are 147 peaks. The algorithm divides the data into 148 rising phases and 147 falling phases in VIX. Duration of falling volatility-phases tends to be longer than rising volatility-phases, with an average and median of 18.50 and 16.00 for rising phases and 26.00 and 20 for falling phases. The minimum duration contains of one day for rising volatility-phases and three days for falling volatility-phases. This means that the change in VIX increased more than 30 percent in one day only and decreased over 30 percent over three days. The maximum duration is 72 days and 101 days for rising and falling volatility-phases respectively. The duration is quite stable over the half samples of the period with longer duration for falling volatility-phases. On the other hand, in terms of amplitude the rising volatility-phases tends to increase more than the falling phases decreases. The average amplitude is 53.25 and 31.54 for rising and falling phases respectively. The maximum rising amplitude is 325.62 and are the percentage increase over one phase, in fact this maximum phase is from 2008.25.08 to 2008.27.10 when the financial crisis really hit. All over the amplitude of rising and falling volatility-phases are stable over the half sample.

Statistics	1990-2015		1990-2002		2003-2015	
	Rising	Falling	Rising	Falling	Rising	Falling
Number of phases	148.0000	147.0000	72.0000	71.00000	76.0000	75.0000
Minimum duration	1.0000	3.0000	1.0000	4.0000	2.0000	3.0000
Average duration	18.4932	25.9524	20.0694	25.8310	17.4211	25.9733
Median duration	16.0000	20.0000	17.0000	20.0000	15.5000	20.0000
Maximum duration	72.0000	101.0000	72.0000	92.0000	46.0000	101.0000
Minimum amplitude	6.6718	9.9274	10.4505	9.9274	6.6718	10.2672
Average amplitude	52.2530	31.5431	47.4483	29.4518	57.1839	33.6524
Median amplitude	41.7870	30.6833	37.6676	27.4027	43.9413	34.2083
Maximum amplitude	325.6247	60.5642	181.8238	58.8398	325.6247	60.5642
Average cum. return	-7.1054	8.8428	-3.0113	3.9058	-4.1222	4.9325
Mean daily return, μ	-0.2596	0.2318	-0.2084	0.2130	-0.3113	0.2532
p-value of $H_0: \mu = 0$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Daily std. deviation	1.2152	1.0277	1.1265	0.9477	1.3028	1.0924
Skewness	-0.7052	0.5259	-0.4532	0.5622	-0.8264	0.4883
Kurtosis	9.1650	7.7525	3.9561	2.8308	11.7496	10.2905
Shapiro-Wilk test, p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Autocorrelation, ρ_1	-0.0060	-0.0480	0.0230	0.0270	-0.0400	-0.1010
p-value of $H_0: \rho_1 = 0$	0.7337	0.0030	0.3860	0.2541	0.1473	0.0009

Table 6-1 Summary statistics of periods of rising and falling VIX using algorithm by BB. Duration is measured in days. Amplitude, average cumulative return, mean daily return and daily standard deviation are measured in percentages. The estimation of the mean daily returns and the daily standard deviation.

The average cumulative return and the mean daily return is negative in phases of rising volatility, while positive in phases of falling volatility. This is the case for the full period, as well as the first half and the second half of the period. The skewness is negative for all the rising phases, and positive for all the falling phases. This means that the tail is right sided for the rising phases and left sided for the falling phases. Kurtosis is higher for the rising phases than it is for the falling phases of all periods. A normal distribution has a skewness of 0 and a kurtosis of 3. From our data the sample is not normally distributed, the first half of the period is closest to a normal distribution. The Shapiro-Wilk test also reject the null hypothesis at a 99% confidence interval. This means that the data sample is not normally distributed. We cannot reject the null hypothesis about first-order autocorrelation at a 95% confidence interval for the rising phase of the full period, both phases of the first period, and the rising phase of the second half of the period, this means that these are uncorrelated. We reject the null hypothesis about autocorrelation for the falling phases for both the full and second period, this means the disturbance term is correlated with other disturbance. The autocorrelation is negative which means that the previous value was positive and the current value is negative.

6.2.2 Algorithm by Lunde and Timmermann

The peaks found by the Algorithm by LT are 127. The algorithm divides the data into 127 rising phases and 126 falling phases in VIX. With LT algorithm the phases are less than BB algorithm due to the change in relative VIX have to be 30 or 20 percent for rising and falling phases. The average duration is longer during falling volatility-phases than rising volatility-phases and are stable over half of the sample period. The pattern is equal the results found by the algorithm by BB for amplitude as well as the change in VIX has higher percentage for rising phases than for falling phases. The rising phases have an average duration of 19.63 and an average amplitude of 59.01. For the falling phases the average duration is 31.96 and average amplitude is 34.92.

Statistics	1990-2015		1990-2002		2003-2015	
	Rising	Falling	Rising	Falling	Rising	Falling
Number of phases	127.0000	126.0000	58.0000	57.0000	69.0000	69.0000
Minimum duration	1.0000	3.0000	3.0000	3.0000	1.0000	3.0000
Average duration	19.6299	31.9606	23.3275	33.7894	16.8116	30.6087
Median duration	16.0000	23.0000	17.5000	26.0000	15.0000	19.0000
Maximum duration	161.0000	248.0000	161.0000	130.0000	57.0000	248.0000
Minimum amplitude	30.0392	20.0480	30.5052	20.0480	30.0392	20.3531
Average amplitude	59.0104	34.9235	56.4841	34.0231	61.7499	35.8799
Median amplitude	46.4186	34.0834	45.0809	33.5304	46.9139	34.7229
Maximum amplitude	271.8767	60.7027	179.0511	60.2194	271.8767	60.7028
Average cum. return	-7.2935	9.0309	-3.0416	3.9360	-4.2455	5.0558
Mean daily return, μ	-0.2925	0.2225	-0.2248	0.2044	-0.3660	0.2394
p-value of $H_0: \mu = 0$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Daily std. deviation	1.1942	1.0517	1.1208	0.9609	1.2714	1.1251
Skewness	-1.0168	0.5982	-0.5414	0.5489	-1.3309	0.6045
Kurtosis	6.3556	10.8356	4.1210	2.7178	7.4335	14.4749
Shapiro-Wilk test, p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Autocorrelation, ρ_1	0.0260	-0.0380	0.0390	0.0300	0.0040	-0.0830
p-value of $H_0: \rho_1 = 0$	0.1883	0.0143	0.1528	0.1935	0.8916	0.0001

Table 6-2 Summary statistics of periods of rising and falling VIX using algorithm by LT.

Duration is measured in days. Amplitude, average cumulative return, mean daily return and daily standard deviation are measured in percentages. The estimation of the mean daily returns and the daily standard deviation.

The last part of the summary statistics in algorithm by LT is approximate the same as in the BB algorithm. The same statements and conclusions are made.

7 Event Study of Abnormal Returns

In this thesis the event study methodology is used to test abnormal returns around VIX volatility peaks determine by rising and falling phases.

7.1 Methodology

To analyze the phenomenon, we used a parametric two-sample test as well as two nonparametric tests to test for the mean abnormal return using 5, 10, 15, and 20 days before and after a peak in VIX. The choice of using both tests is because when using a parametric test, the residuals have to be normally distributed; this is not an assumption when using nonparametric tests. Then we used the event study methodology with dummy variables to test for the abnormal returns around the VIX volatility peaks, we applied the procedures of the algorithm by BB and the algorithm by LT to identify peaks.

7.1.1 The Welch Two Sample t-test

The Welch two sample t-test is a method calculated by default. When using a parametric two sample t-test we are testing two different groups and assuming that both groups are normally distributed with $N(\mu_x, \sigma_x^2)$ and $N(\mu_y, \sigma_y^2)$. The hypothesis will be expressed as,

$$H_0: \mu_x = \mu_y,$$

$$H_1: \mu_x \neq \mu_y,$$

where the null hypothesis states that the two samples are equal, whereas the alternative hypothesis states that the two samples are different. The test statistics can be expressed as follows,

$$t = \frac{\mu_x - \mu_y}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}}, \quad (7.1)$$

where n_x and n_y represents the sample size, μ_x and μ_y the sample mean, s_x^2 and s_y^2 represents the standard deviation or the sample variance of the two groups. To make the conclusion about rejecting the null hypothesis, the following equation must hold,

$$|t| > t_{1-\alpha/2,v}, \quad (7.2)$$

where $t_{1-\alpha/2,v}$ is the critical value of the t-distribution, and v represents the degrees of freedom, and is calculated as (Taeger & Kuhnt, 2014),

$$v = \frac{\left(\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}\right)}{\frac{\left(\frac{s_x^2}{n_x}\right)^2}{n_x-1} + \frac{\left(\frac{s_y^2}{n_y}\right)^2}{n_y-1}}. \quad (7.3)$$

The p-value are calculated using the equation,

$$p = 1 - P(T \leq t). \quad (7.4)$$

Rejecting the null hypothesis with regard to the p-value, the p-value must be less than the chosen level of significance.

7.1.2 Two Sample Wilcoxon Test

The two sample Wilcoxon test is a nonparametric test, which means that we can avoid making assumption about the normal distribution (Dalgard, 2008). This is done by replacing generally obtained data with corresponding order statistics. The method is based on two independent samples, and whether the two samples differ by a shift in location. The two samples are randomly drawn. The Wilcoxon test compares two samples $(x_1, x_2, \dots, x_n$ and $y_1, y_2, \dots, y_n)$ where the observations is given by $N = n_1 + n_1$. We will then compute the sum of the ranks of each of the samples (Tamhane & Dunlop, 2000),

$$w_1 + w_2 = 1 + 2 + \dots + N = \frac{N(N+1)}{2}. \quad (7.5)$$

The hypothesis are expressed as,

$$H_0: x = y,$$

$$H_1: x \neq y,$$

where the null hypothesis state that the two samples are equal. We can reject the null hypothesis if the p-value is lower that the chosen level of significance.

7.1.3 The Kolmogorov-Smirnov Two Sample Test

The Kolmogorov-Smirnov two sample test is a nonparametric test using two different data samples to analyze for independence. The data must contain of two assumptions. First, all the observations for data x_n and y_n are both random samples from continuous populations where all the values of x and y is mutually independent and identically distributed. Second, the two samples must be independent. When these two assumptions hold we can calculate the two sided test statistics Z . The first step is to calculate the empirical distribution of $X_n(t)$ for x and $Y_m(t)$ for y , where t represents the time index.

$$X_n(t) = \frac{\text{number of observed } X_i \leq t}{n}, \quad (7.6)$$

$$Y_m(t) = \frac{\text{number of observed } Y_i \leq t}{m}, \quad (7.7)$$

where n is the sample size of x , and m is the sample size of y . Next step is to calculate the divergence, D , where we compute the absolute values of each of the two groups and subtract. The equation are expressed as,

$$D = |X_n(t) - Y_m(t)|. \quad (7.8)$$

We will use the largest divergence when computing the test statistics. The test statistics are expressed as,

$$Z = D_{max} \sqrt{\frac{nm}{n+m}}. \quad (7.9)$$

Once we have our test statistics we want to compute p-value to determine whether the two samples are statistically significant (Corder & Foreman, 2014). The p-value is calculated according to criteria,

$$\text{if } 0 \leq Z < 0.27,$$

$$\text{then } p = 1 - \frac{2.506628}{Z} (Q + Q^9 + Q^{25}),$$

$$\text{where } Q = e^{-1.233701Z^{-2}},$$

$$\text{if } 1 \leq Z < 3.1,$$

$$\text{then } p = 2(Q - Q^4 - Q^9 - Q^{16}),$$

$$\text{where } Q = e^{-2Z^2}.$$

The hypothesis will be expressed as,

$$H_0: X(t) = Y(t) \text{ for every } t,$$

$$H_1: X(t) \neq Y(t) \text{ for at least one value of } t.$$

The null hypothesis state that there is no difference between the two samples. The alternative hypothesis state that there is a difference between the two samples but do not indicate in which direction. When we have the p-value we can find out if the two samples are statistically significant or not. We will reject the null hypothesis if the p-value is less than the chosen level of significance, α .

7.1.4 Event Study with Dummies

The event study methodology has not changed much since it was introduced by Fama, Fisher, Jensen and Roll (1969). Event studies have a long history and have many applications. The intention is to evaluate the impact of a specific event by measuring the abnormal return (AR) as well as the cumulative abnormal return (CAR). The event study analysis is followed by a general flow and not by a clear structure (MacKinlay, 1997). It is important to have a clear definition about the event of interest and the event period, which include the event date, the event window and the estimation window. The days of interest are the peak dates for the VIX index, and the intention is to analyze relevant data about the event windows around the event dates. The event window is the period studied. In this study the event window contains of 21 trading days, the peak date, 10 days before and 10 days after the peaks. However, more than one event is studied; every peak is one event in itself. The event window consists of the mean

of all event windows. Days in a single window that are located in another phase of rising and falling VIX are eliminated from the mean by dummy variables. To estimate the parameters of the model we use the ordinary least square (OLS) method. This method is also used to measure the abnormal returns and cumulative abnormal returns. When modeling the daily returns both statistical models and economic models can be used. Statistical models follow statistical assumptions and the behavior of asset returns, while the economic model follows economic assumptions and rely on the investors behavior (MacKinlay, 1997). The event date is defined where $\tau = 0$ and it is the main date for the event study. The event window is represented where $\tau \in [-10,10]$, and according to MacKinlay (1997) it should be longer than the specific period of interest. The estimation window is the whole period studied. As a basis for the normal return, estimation period includes all observations from daily return in S&P 500. The mean for all rising and falling periods are used to calculate an overall normal return, μ . To measure abnormal return for day $\tau \in [-10,10]$, the event windows are of 10 days before and 10 after the event dates. Some of the rising and falling periods are shorter than 10 days. To avoid the estimation of the event windows to include data of the opposite period dummy variables for each day are used (Binder, 1998). The return for each day in the event window can be explained as follows,

$$r_t = \mu + \sum_{\tau=-10}^{10} \gamma_{\tau} D_{t,\tau} + \varepsilon_t, \quad (7.10)$$

where μ denotes the normal return, γ_{τ} denotes abnormal return on event day, $D_{t,\tau}$ is a dummy variable with value one on event day, τ , and zero otherwise. The dummy variable makes sure not to include observations from the opposite phase on either side of the peak.

Cumulative Abnormal Returns (CAR) in the event window is the sum of abnormal return, summarized for each day of the event window. On the 10th day after the peak CAR represents the total abnormal return in the window. The formula for CAR is as follows,

$$CAR_{\tau} = \sum_{\tau=-10}^{\tau} \gamma_{\tau}. \quad (7.11)$$

To investigate further whether the mean daily abnormal return over the event window depends on the amplitude of the peak, a sort of the rising amplitude has been done. Afterwards the analysis looks at the abnormal return in the below median part of the amplitude and the above median part of the rising amplitude to test the robustness of our findings. In addition, the rising

amplitude give a good estimation of the height of the peak, but can misplace high peaks to the below median side due to only the one side relative height are used, the rising amplitude. If a trough has a high VIX value this can give a peak a low rising amplitude but still the peak can have one of the largest values of falling amplitude on the other side.

7.2 Empirical Results

7.2.1 Results of Mean Abnormal Return from Algorithm by Bry and Boschan

To find out if the volatility peak means something for the results around the volatility peak, three tests are used to find if the results before the peak are different from the results after the peak. The tests used are the Welch two sample test, the two sample Wilcoxon test and Kolmogorov-Smirnov two sample test. The return N days before and after the peak show the mean abnormal daily return for N days. The mean abnormal daily return is expected to decrease in the days before the peak and increase after the peak.

N	Before	After	p-value		
			Welch test	Wilcoxon test	Kolmogorov-Smirnov test
5	-0.3505	0.4966	0.0000	0.0000	0.0000
10	-0.2220	0.3088	0.0000	0.0000	0.0000
15	-0.1694	0.2217	0.0000	0.0000	0.0000
20	-0.1273	0.1636	0.0000	0.0000	0.0000

Table 7-1 Mean abnormal return around volatility peaks, algorithm by BB. Mean Abnormal Returns is represented in percentage. N represents days before and after the peak. Before represent the mean abnormal return before all peaks at N days. After represent the mean abnormal return after all peaks at N days.

From table 7.1 we can see that mean daily abnormal return for N (5, 10, 15 and 20) days before the peak is negative and after the peak the mean daily abnormal return is positive. This result is what we expected. The entire test represented in table 7.1 has a p-value lower than 0.01 for all N days. This means that we can reject the null hypothesis, and all the three tests are statistically significant at a 1% level. The Welch Two Sample t-test states that the daily mean abnormal return before the peak and the daily mean abnormal return after the peak are different from each other. In this test one assumption is that the data should be normally distributed, which is not the case in our data sample. We can therefore not rely on this result. We will therefore make our conclusions based on the two other tests which are not parametric tests but nonparametric tests. The two sample Wilcoxon test states that the two samples, the daily mean abnormal return before the peak and the daily mean abnormal return after the peak, are not

equal; there exist a difference. The Kolmogorov-Smirnov two sample test give a p-value that will be approximate in the presence of ties. This test also states that there is a difference between the daily mean abnormal return before the peak and the daily mean abnormal return after the peak, but it does not indicate in which direction. Both tests conclude that there is a difference between the mean daily abnormal return before the peak and the mean daily abnormal return after the peak.

7.2.2 Event Study of Abnormal Returns from Algorithm by Bry and Boschan

The results of the estimation of the regression (7.10) is reported in table 7.2. It shows clear evidence of abnormal return around VIX volatility peak. The minimum value in the event window for all peaks is at the peak day of -1.4090%. The mean abnormal return has negative estimates before and including the peak day, where it has 5 days that are significantly different from zero. Right after the peak the mean abnormal return has its maximum value within the event window of 0.8790%, the estimates after the peak day is positive and seven days have significant results. The mean cumulative abnormal return is negative for all days in the event window where the minimum estimate is at the peak day of -4.1060%. The mean cumulative abnormal return decrease toward and including the peak day while it is increasing towards zero right after the peak day. The mean abnormal cumulative return is statistically significant different from zero at the end of the event window while the mean abnormal return is insignificantly different from zero. The pattern is the same in the two sub periods, which shows the first half and the second half of the total period. The table represents the mean daily abnormal return and the mean cumulative abnormal return over the event window. Table 7.2 reports the estimates, and the bold text represents the estimates that are statistically significant at a 5% level.

N	All peaks		1st sub period		2nd sub period		Above median		Below median	
	AR	CAR	AR	CAR	AR	CAR	AR	CAR	AR	CAR
-10	-0.1396	-0.1396	-0.2674	-0.2674	0.0066	0.0066	-0.2286	-0.2286	-0.0386	-0.0386
-9	-0.1160	-0.2556	-0.0485	-0.3159	-0.2245	-0.2179	-0.2436	-0.4722	-0.0173	-0.0559
-8	-0.2010	-0.4566	-0.0095	-0.3254	-0.3407	-0.5585	-0.3410	-0.8131	-0.0600	-0.1159
-7	-0.0595	-0.5162	0.0973	-0.2281	-0.2181	-0.7766	-0.2979	-1.1110	-0.1911	0.0751
-6	-0.3174	-0.8335	-0.3146	-0.5428	-0.3351	-1.1117	-0.4703	-1.5814	-0.1499	-0.0747
-5	0.0140	-0.8195	-0.0825	-0.6253	0.1233	-0.9884	-0.0017	-1.5830	-0.0267	-0.0481
-4	-0.2462	-1.0657	-0.0361	-0.6614	-0.4178	-1.4062	-0.4650	-2.0480	-0.0286	-0.0766
-3	-0.4600	-1.5257	-0.4337	-1.0951	-0.4849	-1.8911	-0.7734	-2.8213	-0.1518	-0.2284
-2	-0.4394	-1.9651	-0.4927	-1.5878	-0.3807	-2.2717	-0.5340	-3.3554	-0.3387	-0.5671
-1	-0.7319	-2.6970	-0.5511	-2.1390	-0.8510	-3.1227	-1.0873	-4.4427	-0.3614	-0.9285
0	-1.4090	-4.1060	-1.2840	-3.4230	-1.5102	-4.6329	-1.7907	-6.2334	-1.0450	-1.9735
1	0.8790	-3.2270	0.6211	-2.8019	1.0983	-3.5346	1.4005	-4.8329	0.3782	-1.5954
2	0.5645	-2.6625	0.5389	-2.2630	0.6161	-2.9184	0.5656	-4.2673	0.5780	-1.0174
3	0.4151	2.2475	0.5103	-1.7527	0.3934	-2.5250	0.4956	-3.7717	0.3216	-0.6958
4	0.3109	-1.9366	0.2929	-1.4598	0.3247	-2.2003	0.4992	-3.2725	0.1232	-0.5726
5	0.3425	-1.5941	0.3402	-1.1196	0.2991	-1.9012	0.3789	-2.8936	0.3059	-0.2667
6	0.1741	-1.4200	0.1226	-0.9969	0.2738	-1.6274	0.1685	-2.7250	0.1787	-0.0880
7	0.1515	-1.2685	0.1784	-0.8186	0.1284	-1.4990	0.1621	-2.5630	0.1330	0.0450
8	0.1798	-1.0887	0.1263	-0.6923	0.2679	-1.2311	0.3172	-2.2458	0.0733	0.1183
9	-0.0309	-1.1196	0.0708	-0.6215	-0.1556	-1.3867	-0.1879	-2.4338	0.0831	0.2014
10	0.1340	-0.9856	0.1467	-0.4748	0.1701	-1.2166	0.3628	-2.0710	-0.0479	0.1535

Table 7-2 Mean abnormal return over the event window, algorithm by BB.

AR denotes the mean abnormal return and CAR denotes the mean cumulative abnormal return. All returns are represented in percentages. The bold text indicates returns that are statistically significant at the 5% level. Above median denotes the results of peaks with amplitude above median while the below median denotes the results of peaks with amplitude below median.

The results for all peaks show evidence of abnormal returns around VIX volatility peak. We can reject the null hypothesis of zero mean daily abnormal return of 5 days including the peak date, and 7 days following the peak date. The mean daily abnormal return has a clear pattern where it is negative and decreasing before and until the peak day, and positive after the peak. The mean cumulative abnormal return shows the same pattern, thus the mean cumulative abnormal return is statistically different from zero of 9 days including the peak day, and 10 days following the peak day. The mean cumulative abnormal return is negative and decreasing including the peak day, right after the peak day the mean cumulative abnormal return increases towards zero, but remains negative.

When looking at the two sub periods, the mean abnormal return in the first half has 12 of 21 days that are significantly different from zero, where 6 days including the peak day, and 6 days following the peak day. While the second half has 14 of 21 days that are significantly different from zero, where 7 days including the peak day, and 7 days following the peak day. The mean abnormal return in both sub periods has its minimum value at the peak day (-1.2840% and -1.5102%). Both sub periods have also positive and decreasing values before and including the peak day, while positive and increasing values after the peak day.

At the start and at the end of the event window most of the values for the mean abnormal return are insignificantly different from zero. The last day in the event window the cumulative abnormal return is statistically significant different from zero, as well as the days before. This

is also the case when looking at the two sub periods, where in the first half the last days of the event window is statistically significant different from zero and this is the case for the second half of the period. The results from the event study of abnormal return are presented graphically in the figures that follows.

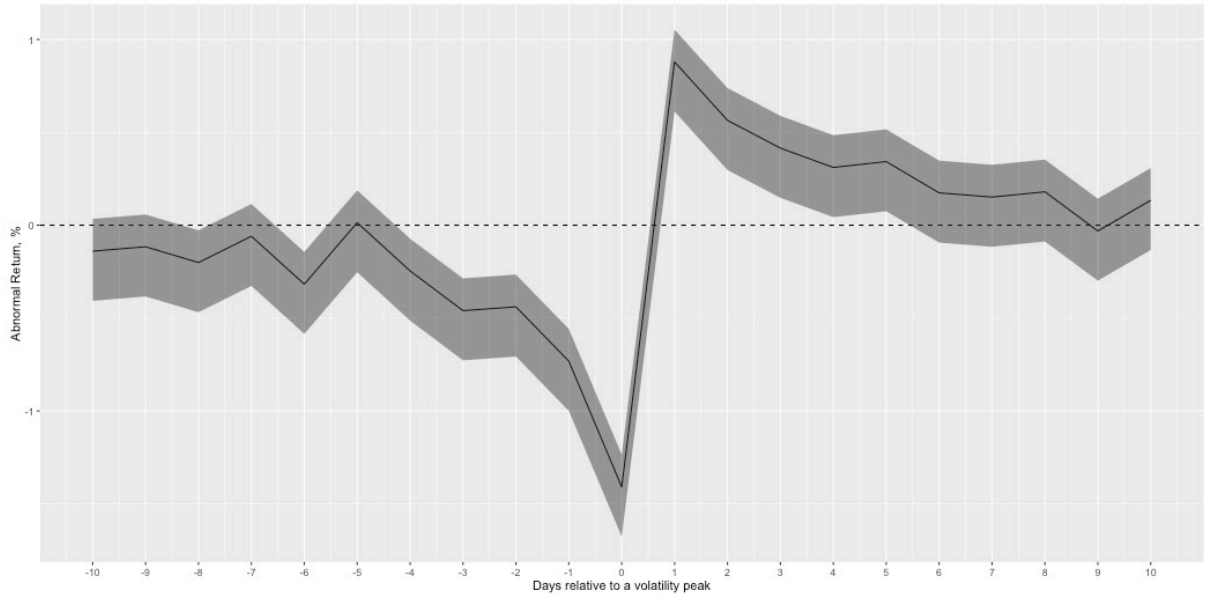


Figure 7-1 Mean daily abnormal return over the event window, algorithm by BB.

The shaded area indicates the 95% confidence interval.

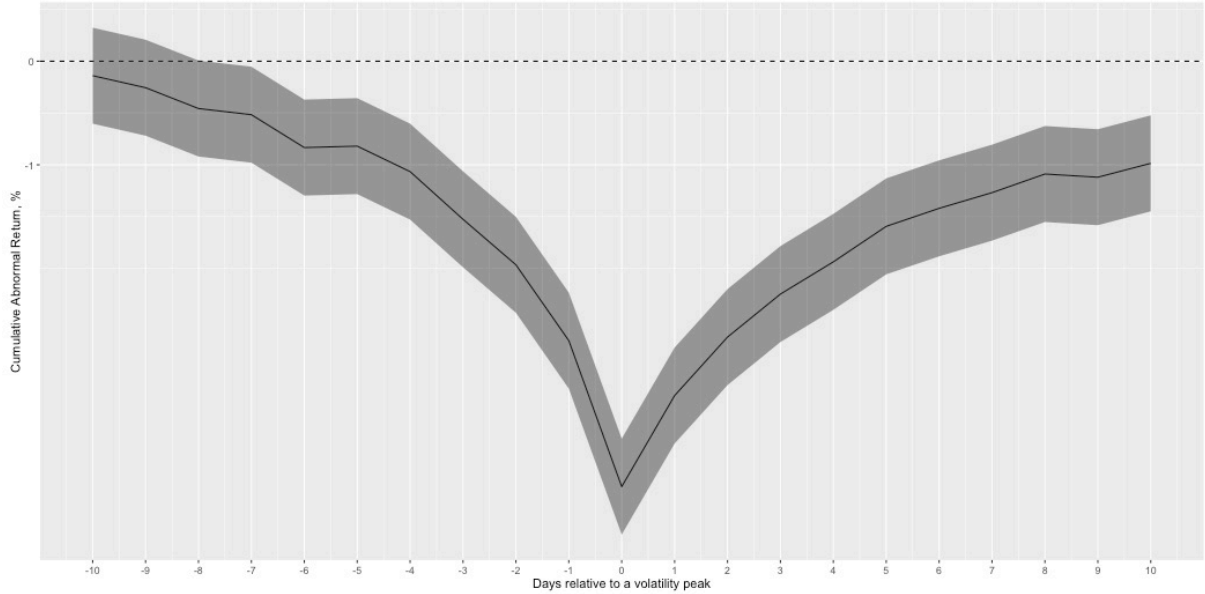


Figure 7-2 Mean cumulative abnormal return over the event window, algorithm by BB..

The shaded area indicates the 95% confidence interval.

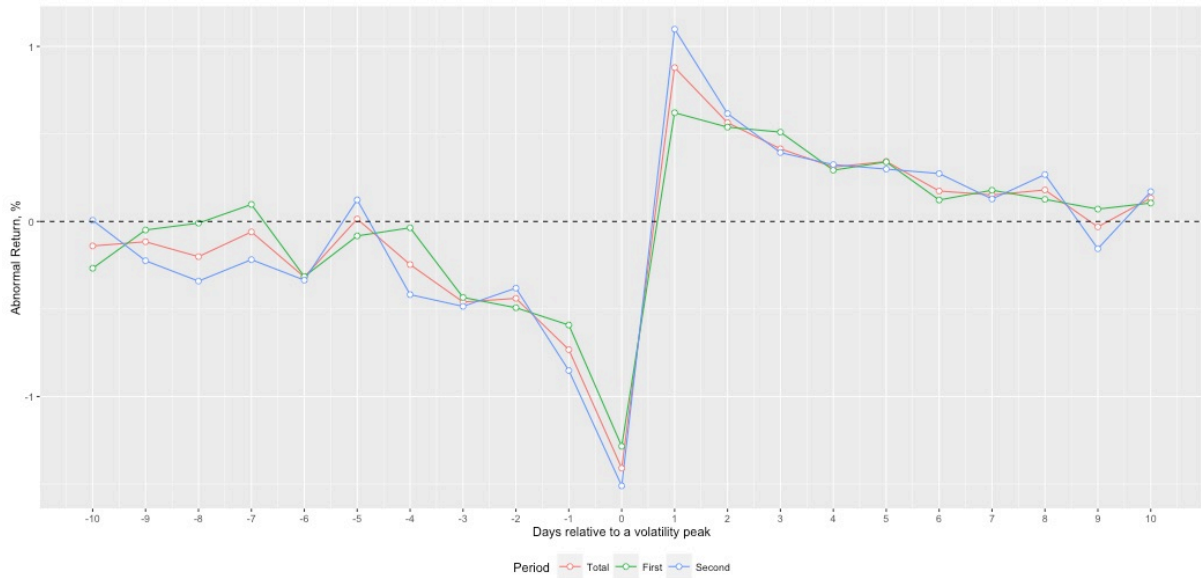


Figure 7-3 Mean abnormal return for over the event window for full period, first half and second half of the sample, algorithm by BB.

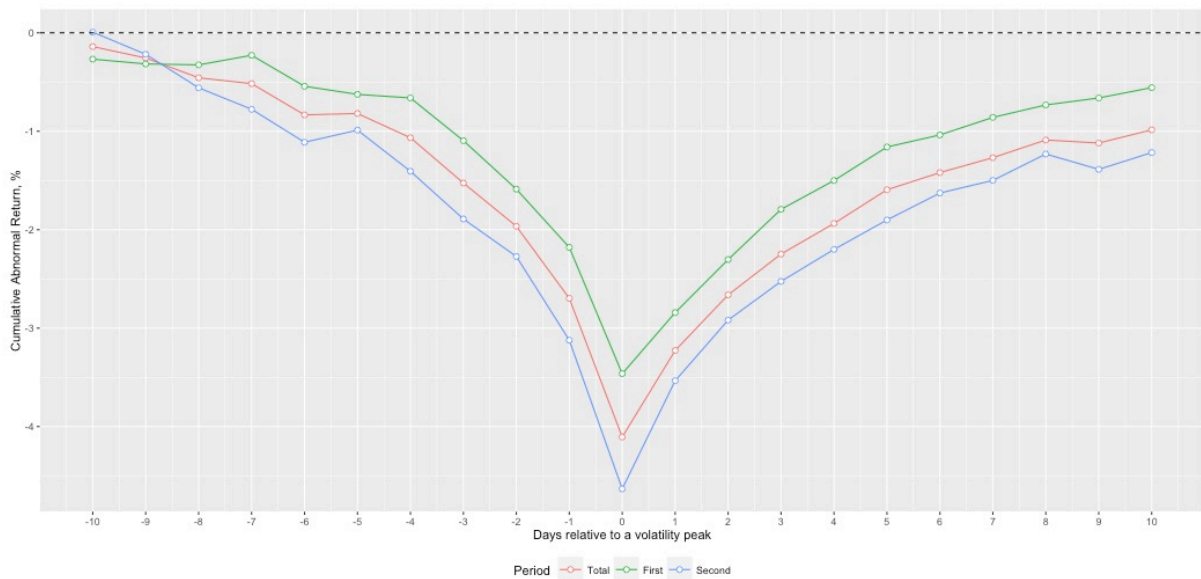


Figure 7-4 Cumulative abnormal return for over the event window for full period, first half and second half of the sample, algorithm by BB.

7.2.3 Robustness Test from Algorithm by Bry and Boschan

The results of the above median amplitude and the below median amplitude is represented in table 7.2. The results are robust and quite stable in both the above median amplitude and the below median amplitude. The above median amplitude has 13 days that are statistically significant different from zero which is more than the below median amplitude of 11 days. The mean cumulative abnormal return has all estimates after and including the peak day statistically

significant different from zero, this is the case for both above median amplitude and below median amplitude. We can say that there is clear evidence of abnormal returns which depends on the amplitude of a VIX volatility peak.

At the peak day the abnormal return is negative and attains its minimum of -1.7907% for the above median amplitude, while for the below median amplitude the abnormal return is -1.0450%. The abnormal return for both above median amplitude and below median amplitude, shows evidence of more negative abnormal return the larger the amplitude of the peak is before and until the peak day, and more positive after the peak. Thus, it is decreasing before and including the peak day, and increasing towards zero after the peak day.

When looking at the cumulative abnormal return it has negative estimates for almost all days in the event window. The abnormal return has, as stated earlier negative estimates before and including the peak day, while positive estimates after the peak day. This is not the case for the cumulative abnormal returns. From table 7.2 we can see that greater the amplitude of a peak, the more negative is the cumulative abnormal return at the end of the event window. The above median amplitude attains its minimum at the peak day of -6.2334%, at the end it attains -2.0710%, for the below median amplitude at the peak day attains -1.9735% and at the end it attains 0.1535%. The above and below median amplitude over the event window are presented graphically in the figures that follows.

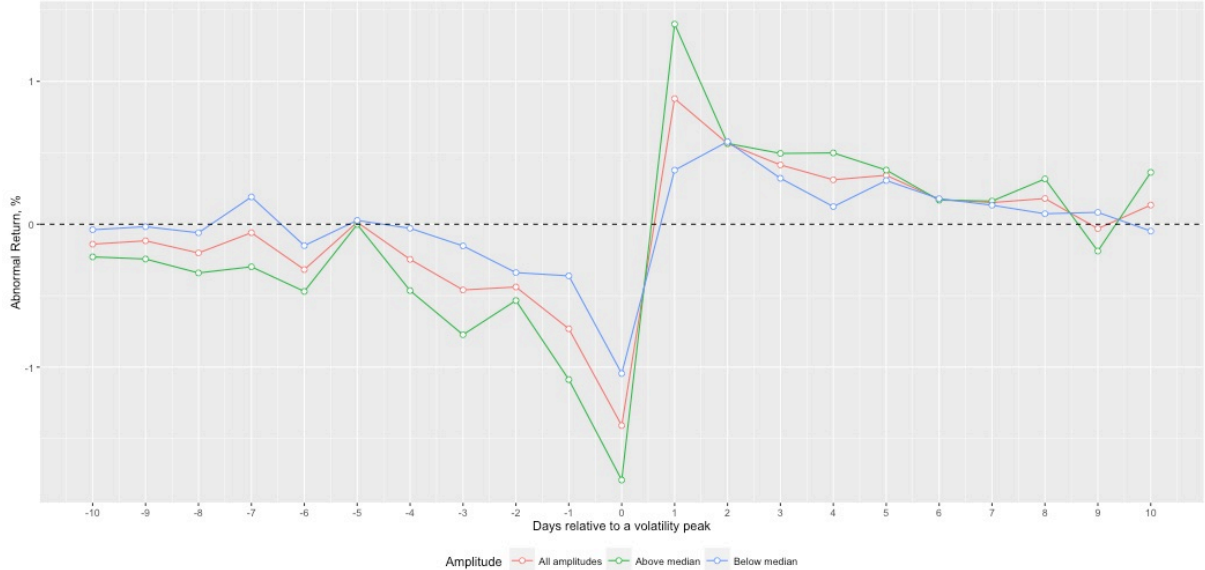


Figure 7-5 Mean abnormal return for over the event window for all peaks, above and below median amplitude of peaks, algorithm by BB.

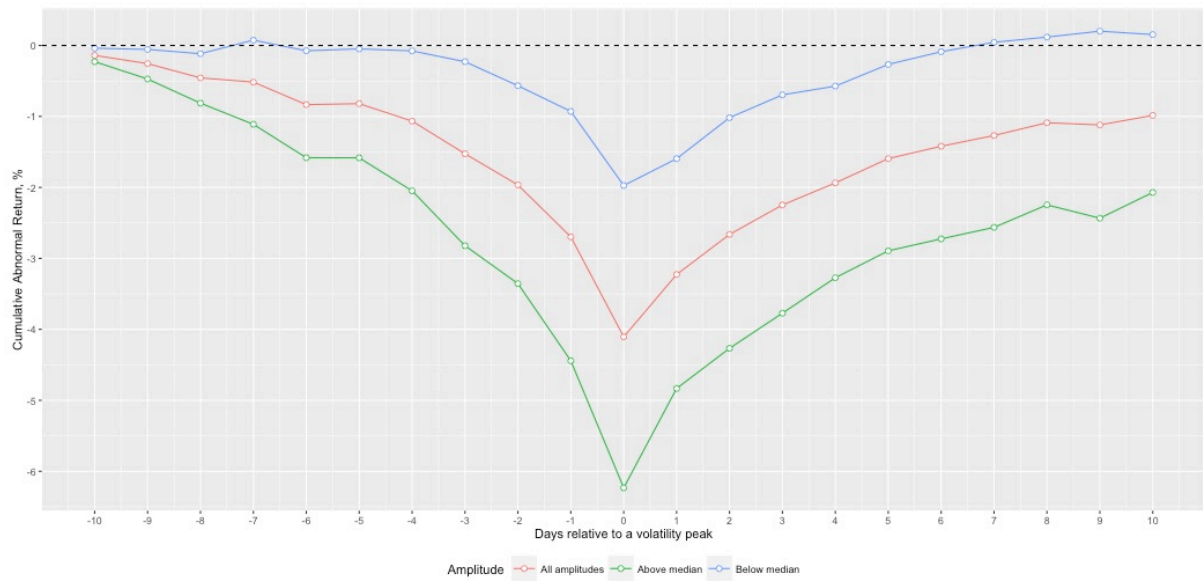


Figure 7-6 Cumulative abnormal return for over the event window for all peaks, above and below median amplitude of peaks, algorithm by BB.

7.2.4 Results of Mean Abnormal Return from Algorithm by Lunde and Timmermann

N	Before	After	p-value		
			Welch test	Wilcoxon test	Kolmogorov-Smirnov test
5	-0.4653	0.5497	0.0000	0.0000	0.0000
10	-0.3132	0.3040	0.0000	0.0000	0.0000
15	-0.2434	0.2305	0.0000	0.0000	0.0000
20	-0.1883	0.1685	0.0000	0.0000	0.0000

Table 7-3 Mean abnormal return around volatility peaks, algorithm by LT. Mean Abnormal Returns is represented in percent. N represents days before and after the peak. Before represent the mean abnormal return before all peaks at N days. After represent the mean abnormal return after all peaks at N days.

The values of the mean daily abnormal return before the peak as well as the mean daily abnormal return after the peak is a bit different from what was represented using the BB algorithm. These differences will not have too much impact because the conclusions will be the same as stated in the BB algorithm. The mean daily abnormal return before the peak is negative for all N days, and the mean daily abnormal return after the peak is positive for all N days, same as in the algorithm by BB. The conclusions for the three two sample tests will be the same as the BB algorithm, thus we will reject the null hypothesis and support the alternative hypothesis.

7.2.5 Event Study of Abnormal Returns from Algorithm by Lunde and Timmermann

The results of the estimation of regression (7.10) using the algorithm by LT are reported in table 7.4. This table show same evidence as in the BB algorithm, it is clear evidence of abnormal return around VIX volatility peaks. The table 7.4 reports only the estimates and the bold text represents the estimates that are significantly significant at a 5% level.

We can reject the null hypothesis of zero mean daily abnormal return of 6 days (including the peak day) preceding the peak day, and 8 days following the peak day. The same number of days as in the BB algorithm. The mean daily abnormal return shows the same pattern as in the BB algorithm, the estimates is positive and decreasing in value before and including the peak day, where it attains its minimum of -1.6207%. The day right after the peak day the mean daily abnormal return attains its maximum of 1.1744%, and after the peak day the mean daily abnormal return is positive and increasing towards zero. When looking at the mean cumulative abnormal return all estimated is negative. The mean cumulative abnormal return is statistically different from zero of 10 days including the peak day, and 10 days following the peak day. The mean cumulative abnormal return is negative and decreasing before and until the peak day where it attains its minimum of -5.7671%, the day after the peak day the mean cumulative abnormal return is still negative but increasing towards zero.

N	All peaks		1st sub period		2nd sub period		Above median		Below median	
	AR	CAR	AR	CAR	AR	CAR	AR	CAR	AR	CAR
-10	-0.3018	-0.3018	-0.3667	-0.3667	-0.2390	-0.2390	-0.3702	-0.3702	-0.2417	-0.2417
-9	-0.4191	-0.7209	-0.3127	-0.6793	-0.5293	-0.7683	-0.6661	-1.0363	-0.1211	-0.3629
-8	-0.2069	-0.9277	-0.1877	-0.8670	-0.1817	-0.9500	-0.2898	-1.3261	-0.1612	-0.5240
-7	-0.2599	-1.1876	-0.0970	-0.9641	-0.4619	-1.4119	-0.5052	-1.8313	-0.0085	-0.5325
-6	-0.3204	-1.5080	-0.2216	-1.1856	-0.4113	-1.8233	-0.4292	-2.2604	-0.1402	-0.6727
-5	-0.0903	-1.5983	-0.1667	-1.3523	-0.0185	-1.8417	-0.0392	-2.2996	-0.1629	-0.8356
-4	-0.4113	-2.0096	-0.1697	-1.5221	-0.6059	-2.4477	-0.6409	-2.9405	-0.1571	-0.9927
-3	-0.6595	-2.6690	-0.6622	-2.1843	-0.6605	-3.1082	-0.9247	-3.8652	-0.3543	-1.3470
-2	-0.4623	-3.1313	-0.4676	-2.6519	-0.4447	-3.5529	-0.4548	-4.3200	-0.4763	-1.8233
-1	-1.0150	-4.1463	-0.8756	-3.5275	-1.1565	-4.7094	-1.1920	-5.5121	-0.8453	-2.6686
0	-1.6207	-5.7671	-1.6209	-5.1484	-1.6101	-6.3195	-1.8878	-7.3999	-1.3681	-4.0367
1	1.1744	-4.5926	0.9565	-4.1919	1.3529	-4.9667	1.6170	-5.7829	0.7212	-3.3154
2	0.6064	-3.9863	0.6098	-3.5821	0.6130	-4.3537	0.4795	-5.3034	0.7446	-2.5708
3	0.3794	-3.6070	0.6441	-2.9380	0.1730	-4.1807	0.5146	-4.7888	0.2402	-2.3306
4	0.3592	-3.2478	0.3226	-2.6154	0.4176	-3.7631	0.5514	-4.2374	0.1686	-2.1620
5	0.2987	-2.9491	0.3018	-2.3136	0.2659	-3.4971	0.2377	-3.9997	0.3776	-1.7844
6	0.2638	-2.6853	0.2448	-2.0688	0.2991	-3.1980	0.3049	-3.6948	0.2051	-1.5794
7	0.2476	-2.4377	0.2656	-1.8032	0.2594	-2.9386	0.2043	-3.4905	0.2959	-1.2835
8	0.3281	-2.1096	0.2557	-1.5475	0.4198	-2.5188	0.4323	-3.0581	0.2412	-1.0422
9	-0.0384	-2.1480	0.0758	-1.4717	-0.1697	-2.6885	-0.1320	-3.1901	0.0625	-0.9797
10	0.1178	-2.0302	0.0966	-1.3751	0.1380	-2.5505	0.4037	-2.7864	-0.1240	-1.1037

Table 7-4 Mean abnormal return over the event window, algorithm by LT.

AR denotes the mean abnormal return and CAR denotes the mean cumulative abnormal return. All returns are represented in percentages. The bold text indicates returns that are statistically significant at the 5% level. Above median denotes the results of peaks with amplitude above median while the below median denotes the results of peaks with amplitude below median.

The mean daily abnormal return is insignificantly significant at the beginning and at the end of the event window for all peaks. The mean cumulative abnormal return is statistically significantly different from zero for 20 days including the peak day. It is only one estimate at the beginning of the event window that is insignificantly different from zero for all peaks. The results from table (7-4) are presented graphically in the figures that follows.

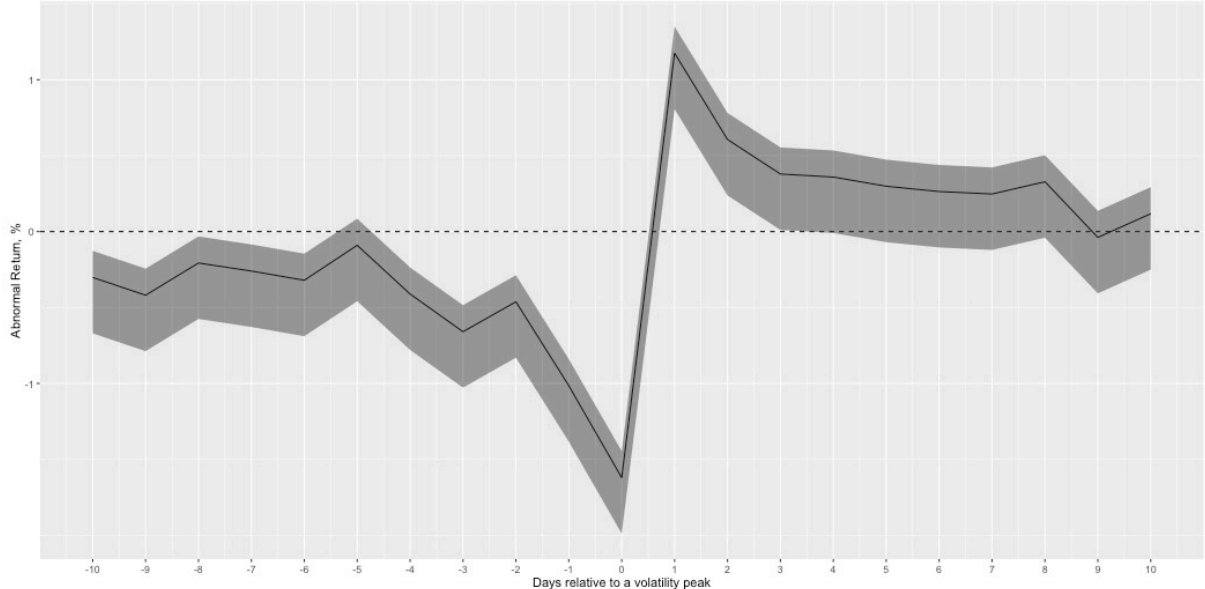


Figure 7-7 Mean daily abnormal return over the event window, algorithm by LT.

The shaded area indicates the 95% confidence interval.

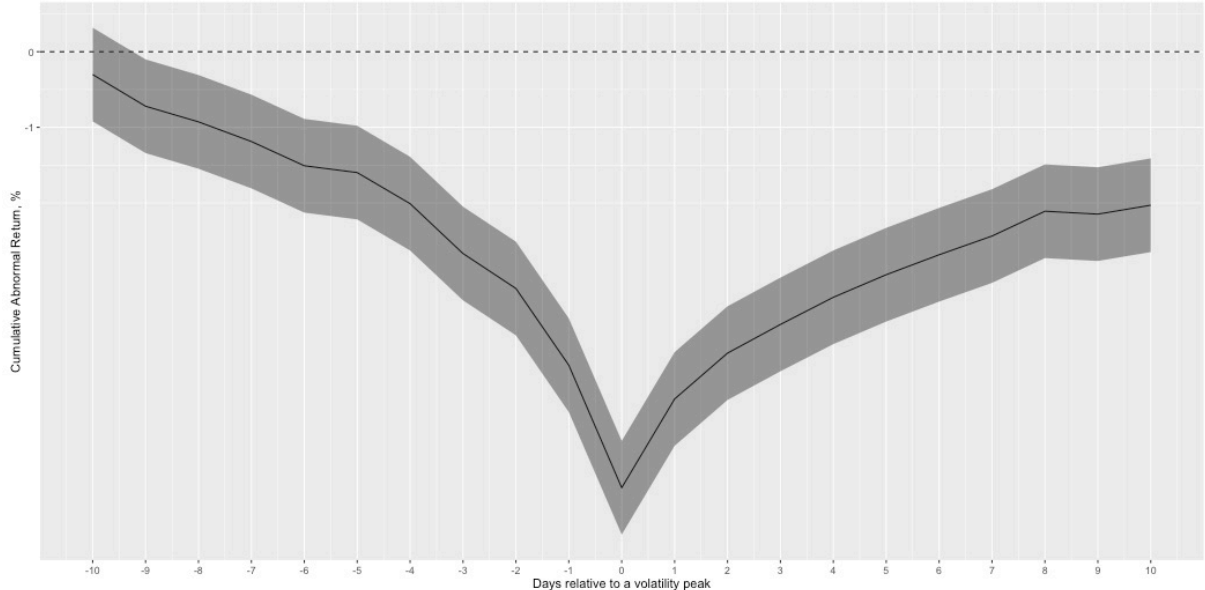


Figure 7-8 Mean cumulative return over the event window, algorithm by LT.

The shaded area indicates the 95% confidence interval.

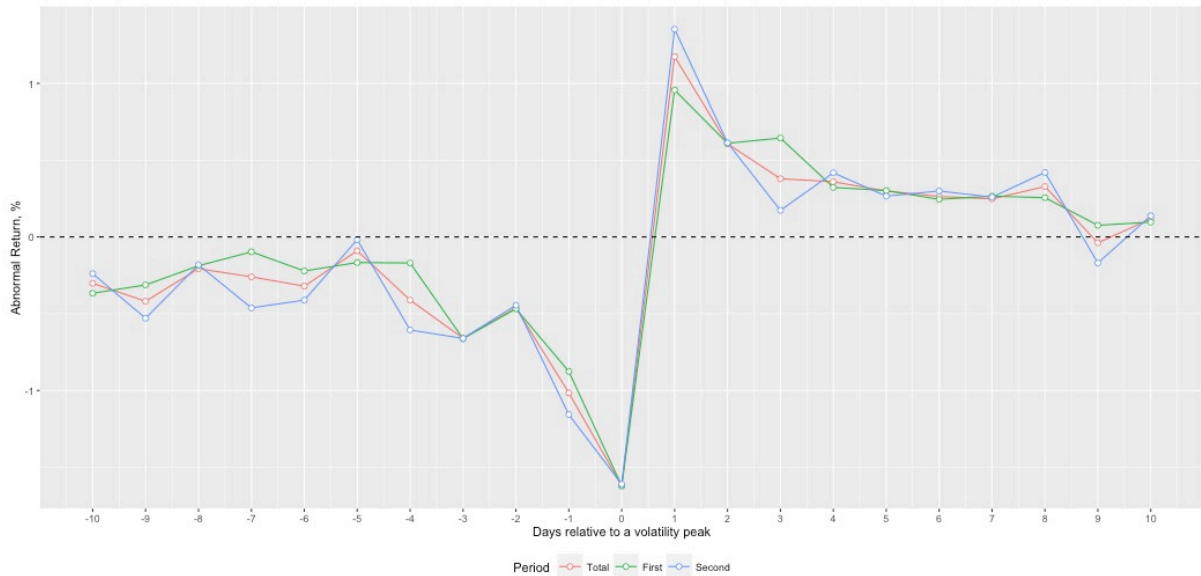


Figure 7-9 Mean abnormal return for over the event window for full period, first half and second half of the sample, algorithm by LT.

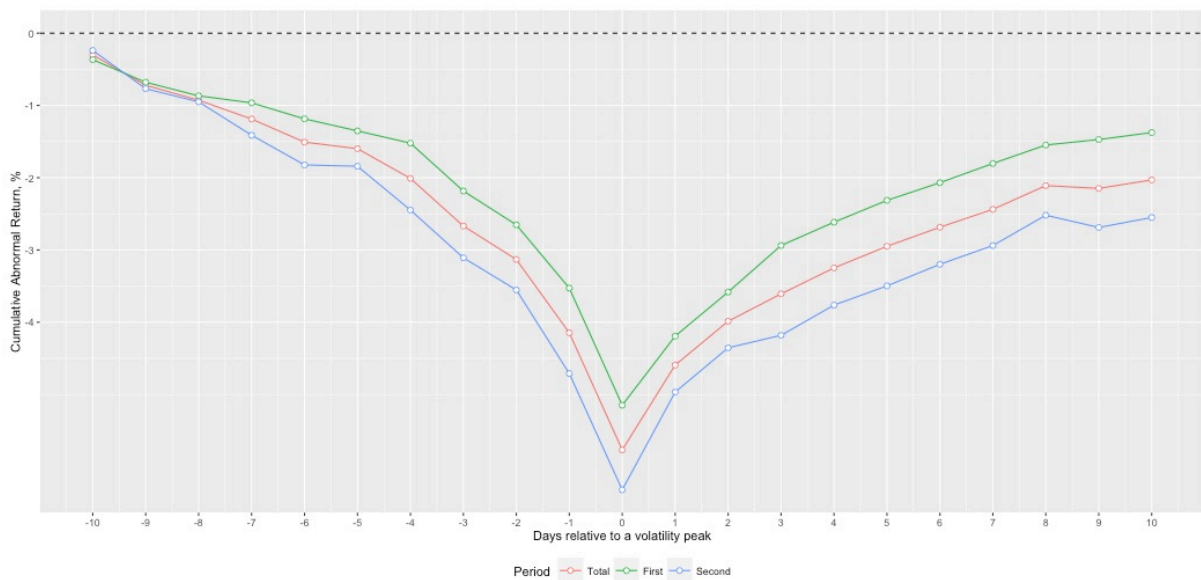


Figure 7-10 Cumulative abnormal return for over the event window for full period, first half and second half of the sample, algorithm by LT.

7.2.6 Robustness Test from Algorithm by Lunde and Timmermann

The results of the above median amplitude and the below median amplitude are represented in table 7.4. The results show robustness of the study. The above median amplitude has 15 days statistically significant different from zero which is more than the below median amplitude of 12 days, which 4 days are before and including the peak and 8 out of 10 days after the peak. For the mean cumulative abnormal return, the significance indicates a clear evidence of abnormal returns which depends on the amplitude of a VIX volatility peak.

At the peak day the abnormal return is negative and have its minimum of -1.8878% on peak day for the above median amplitude, while for the below median amplitude the abnormal return is -1.3681% at peak day. The abnormal return for both above median amplitude and below median amplitude give negative abnormal returns before and including peak day, and positive returns after the peak day. Cumulative abnormal return is decreasing before and including the peak day, and increasing towards zero after the peak day. The above and below median amplitude of the event study are presented graphically in the figures that follows.

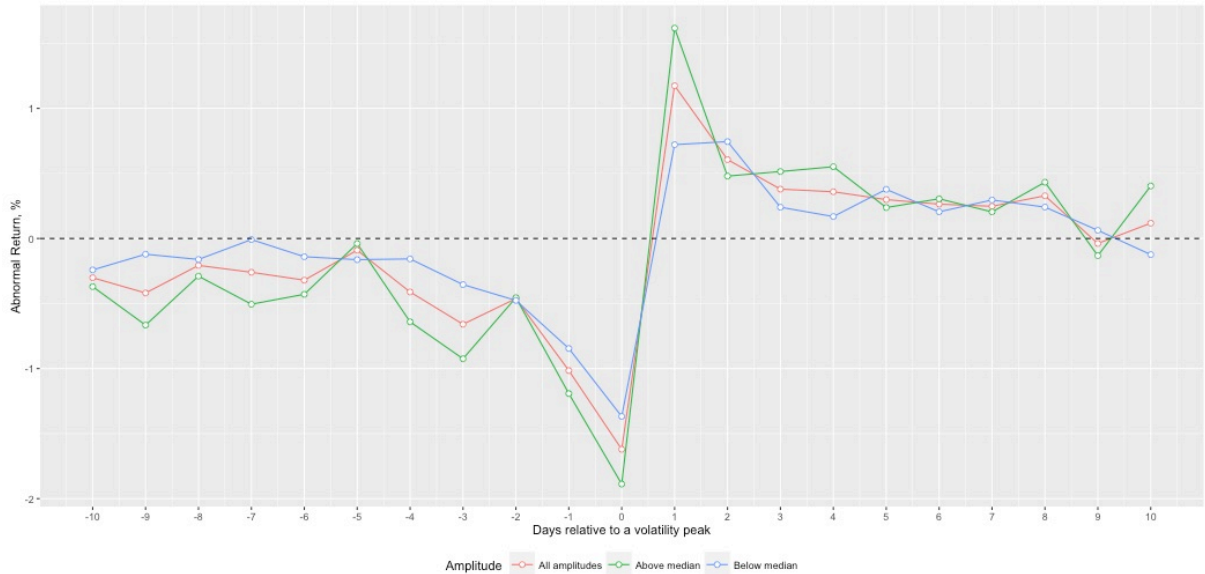


Figure 7-11 Mean abnormal return for over the event window for all peaks, above and below median amplitude of peaks, algorithm by LT.

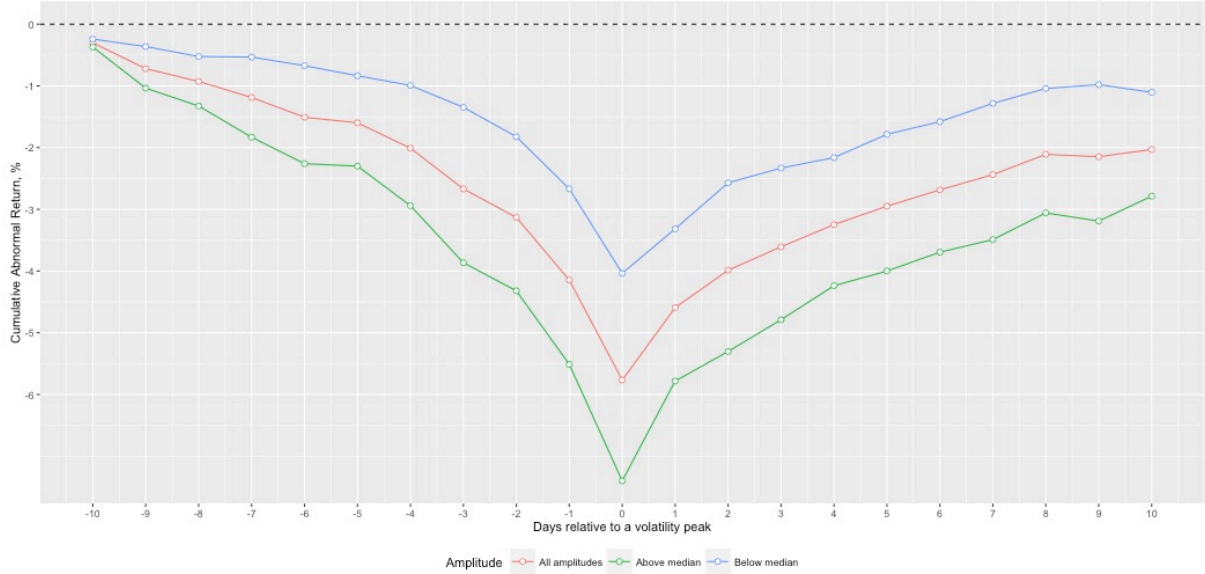


Figure 7-12 Cumulative abnormal return for over the event window for all peaks, above and below median amplitude of peaks, algorithm by LT.

8 Discussion

In this thesis we carried out three different studies. One replica study of Whaley (2000) and Whaley (2009) about the relationship between VIX and S&P 500. A new study of volatility peaks in terms of rising and falling volatility measured by the VIX index. And a new study of abnormal daily returns on volatility peaks in terms of rising and falling volatility measured by the VIX index.

In the study of the relationship between the VIX index and S&P 500 four models were studied: two symmetrical and two asymmetrical regressions. The replica study from Whaley (2000) studied the impact ΔVIX and ΔVIX^+ have on return (asymmetrical), together with the impact ΔVIX has on S&P 500 returns (symmetrical). All findings showed significant result on a 1% level. However, the results from the symmetrical regression (5.2) were more comforting according to the expected outcome. In the symmetrical model the return significantly increased if VIX was equal to zero and VIX had a negative impact on return. On the other hand, the asymmetric regression (5.4), return was not significantly increasing if VIX was equal to zero, the impact of VIX negatively exceeded the symmetrical regression but the variable consisting of positive ΔVIX did not decrease the return, rather increase the return. Whaley's (2000) results from the model is as follows $0.775 - 0.469\Delta VIX_t - 0.238\Delta VIX_t^+$. Even though we did a replica of Whaley (2000) study, we did not reach the expected results in the model. There can be multiple reasons for this. Firstly we use log-returns rather than daily returns. Secondly the time frame as our study includes 16 years of data. In the early measurement of VIX, S&P100 was used in the calculation rather than S&P 500, as well as Whaley (2000) used S&P100 in his study. But these indices give similar return so the difference should be negligible. The third explanation is that Whaley (2000) studied weekly returns by Wednesday-to-Wednesday return and a weekly change in VIX. In our study the increase in VIX may have a delay before the decrease in S&P 500 are in place.

The regression from Whaley (2009) looked at what impact return and all observations with negative return have on VIX (asymmetrical), where the thesis included the impact return in S&P has on VIX (symmetrical). In the regression the results turned out to be as expected, with 99% confidence level, except significant result included the constant as well. Return have a decreasing impact on VIX in both regressions and the negative return exceeds the decreasing change in VIX in the asymmetrical regression (5.5). It is quite strange that the intercept for both

regressions gave significant results, since VIX is not expected to either decrease or increase over time. The explanation can be daily observations. The results from Whaley (2009) are as follows $R_{VIX} = -0.004 - 2.990R_{SPX_t} - 1.503R_{SPX_{t-1}}$. After all, the relationship between S&P 500 and VIX is clear, the variables are negatively correlated where the asymmetrical models gave a bit higher R-squared. But the symmetrical models do give a good prediction of either VIX or S&P 500.

The study of peaks in terms of rising and falling volatility phases is a new study with limited literature to rely on. The study is carried out by two set of algorithms to determine the rising and falling phases, and dating of peaks and troughs, the algorithm by BB and the algorithm by LT. We choose not to use smoothed data as in the study done by Pagan and Sossounov (2003). This because they apply the algorithm by BB in the stock market rather than in the business cycle, and found that not using smoothed data was the best option. In this study we found that both algorithms have a higher average amplitude in raising phases than falling phases, but the average duration is longer during falling phases. The difference between these two algorithm is how they calculate the turning points. By using the algorithm by BB we found 147 peaks, and the algorithm defined 148 rising phases and 147 falling phases. While using the algorithm by LT we calculated 126 peaks, and the algorithm defined 127 rising phases and 126 falling phases. The algorithm by LT consequently has longer average duration in both rising (19.63) and falling (31.96) phases than the algorithm by BB (18.49 and 25.95). The average amplitude is also higher for both rising (59.01) and falling (34.92) phases in the algorithm by LT than the algorithm by BB (52.25 and 31.54), but the algorithm by BB has a higher maximum amplitude for rising (325.62) phases and a lower minimum amplitude for both rising (6.67) and falling (9.93) phases than the algorithm by LT (271.88, 30.04, and 20.05). The study done by Gonzalez, Powell, Shi, and Wilson (2005) find that the bull and bear market is associated with mean return shifts. In our study we found evidence that negative mean daily return is associated with rising phases, and positive mean daily returns are associated with falling phases. Where we see that S&P 500 increases slowly over time until the VIX give a spike peak and the prices drop in S&P 500. Evidence is found in duration, amplitude and the first study in the thesis.

In the study of abnormal returns the purpose was to look at the short term effect of peaks in VIX had on the abnormal return on S&P 500. This is also a new study and we have not found any literature or empirical studies to rely on. We found evidence that there is a difference between the mean daily abnormal return before the peak and the mean abnormal return after

the peak, when using 5, 10, 15, and 20 days. Our results show that the mean daily abnormal returns is negative before the peak, and positive after the peak, this is what we expected. We also found strong evidence of abnormal returns around VIX volatility peaks in the event study. The results of all peaks as well as the first sub period and the second sub period show a clear pattern of the abnormal return and cumulative abnormal return. For the abnormal returns the minimum value of the event window was at the peak date. The estimates were negative before and including the peak date, and positive right after the peak date where it attains its maximum value within the event window. Thus, the estimates decrease towards and including the peak date, while it increases towards zero after the peak date. The estimates of cumulative abnormal return are negative before and after the peak day. The values of the estimates decrease towards and including the peak date, at the peak date the cumulative abnormal return attains its minimum value within the event window. Right after the peak date the values increases towards zero. Almost all the estimates of the cumulative abnormal return are statistically significant at a 5% level only a few days at the beginning of the event window which is not significant. To show the robustness of our results we used the full sample and divided it into above median amplitude and below median amplitude. The results were robust and stable for both above median amplitude and below median amplitude. The abnormal return is statistically significant in days around the peak date, while the cumulative abnormal return is statistically significant some days before the peak date and until the end of the event window.

Behavioral finance theory is more and more relevant, and becoming an important topic. This is because the theoretical models from the classical finance theory are not based on the investors sentiment, and this can have an impact on the results. From the study done by Brown and Cliff (2004) they concluded that the investors sentiment has little or no power for near-term future stock markets returns, while the study by Baker and Wurgler (2006) showed that the investors sentiment has significant cross-sectional effects on stock market prices and returns. In this study it is more interesting to test the impact of investors sentiment on the VIX index, as stated by Whaley (2000) the VIX index is called the “investor fear gauge”. Higher volatility results in higher fear, because the investors set the VIX index and their view is based on the expected future stock market volatility. We have used the behavioral finance theory to make a statement of the relevance of the theory, but we have not done empirical studies of this topic.

A question that rises; is the peak in the VIX an indicator to buy? The most obvious answer is yes, but this is when looking backward. The problem is to locate the peak on the current day.

Furthermore, to understand that, the VIX at some level indicates investors fear (Whaley, 2000). Whaley (2009) says that the VIX index is a forward-looking index because it is measuring volatility that the investors are expected to see, and not the other way around.

8.1 Future Research

In this empirical research we have replicated one study, and created two new studies. Our results are in line with existing studies, but more research is needed on the subject. As we look at the impact of a peak in VIX could be a signal to buy, we could also have looked at if the trough in VIX is a signal to sell a stock, because peaks and trough act opposite of each other in the market. In further research we can include autoregressive models in the first study. In the third study when using the event study methodology, we looked at the short term. To look at the long term on behalf of investors we could have used a wider window and a longer perspective. At last, empirical research about the behavioral finance in relation to S&P 500 and VIX to see if the investors have a significant impact.

9 Conclusion

Even though the asymmetrical relation of S&P 500 and VIX was describe in the literature, the thesis outlined empirical studies to amplify the evidence of the relationship. The evidence is there to conclude with a negative relationship between VIX and S&P 500. By defining the peaks in VIX and analyzing the phases, the statistics clearly states differences between rising and falling volatility-phases. Further, the empirical results from the abnormal return of S&P 500 and what happens in the days around the peaks, gave a clear definition of the significantly negative abnormal return before the peak, and positive abnormal return after the peak. In addition, evidence discussed in this thesis amplify the observed relation of VIX and investors sentiment. The abnormal cumulative return was negative in all periods, including the above and below median amplitude. Since the return was significantly negative before and significantly positive after the peak, the return in total decreased more than increased. We found evidence of abnormal returns around peaks in VIX.

10 References

- Baker, H. K., & Nofsinger, J. (2010). *Behavioral Finance: Investors, Corporations, and Markets*. Hoboken, New Jersey, USA: John Wiley & Sons.
- Baker, M., & Wurgler, J. (2006). Investor Sentiment and the Cross-Section of Stock Returns. *The Journal of Finance*, 61(4), 1645–1680.
- Banerjee, P., Doran, J., & Peterson, D. R. (2007). Implied Volatility and Future Portfolio Returns. *Journal of Banking and Finance*, 31(10), 3183 – 3199.
- Barberis, N., & Thaler, R. (2003). A Survey of Behavioral Finance. In G. M. Constantinides, M. Harris, & R. Stulz, *The Handbook of the Economics of Finance* (pp. 1052-1121). Amsterdam: Elsevier.
- Binder, J. J. (1998). The Event Study Methodology Since 1969. *Review of Quantitative Finance and Accounting*, 11(2), 129-151.
- Bodie, Z., Kane, A., & Marcus, A. J. (2013). *Essentials of Investments* (Ninth ed.). New York: MacGraw-Hill/Irwin.
- Brooks, C. (2008). *Econometrics for Finance* (Second ed.). Cambridge: Cambridge University Press.
- Brown, G. W., & Cliff, M. T. (2004). Investor Sentiment and the Near-Term Stock Market. *Journal of Empirical Finance*, 11(1), 1-27.
- Bry, G., & Boschan, C. (1971). *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs*. New York: National Bureau of Economic Research.
- Cashin, P., McDermott, C. J., & Scott, A. (2002). Booms and Slumps in World Commodity Prices. *Journal of Development Economics*, 69, 277-296.
- Chicago Board Options Exchange. (n.d.). *Chicago Board Options Exchange*. Retrieved March 4, 2016, from VIX: <https://www.cboe.com/micro/vix/vixwhite.pdf>
- Copeland, M. M., & Copeland, T. E. (1999). Market Timing: Style and Size Rotation Using the VIX. *Financial Analysts Journal*, 55(2), 73-81.
- Corder, G. W., & Foreman, D. I. (2014). *Nonparametric Statistics: A Step-by-Step Approach* (Second ed.). New Jersey: John Wiley & Sons.
- Dalgaard, P. (2008). *Introductory Statistics with R* (Second ed.). New York: Springer Science & Business Media.
- Fama, E. F., & Malkiel, B. G. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance*, 25(2), 383-417.

- Fama, E. F., Fisher, L., Jensen, M. C., & Roll, R. (1969). The Adjustment of Stock Prices to New Information. *International Economic Review*, 10(1), 1-21.
- Fernandes, M., Medeiros, M. C., & Scharth, M. (2014). Modeling and Predicting the CBOE Market Volatility Index. *Journal of Banking & Finance*, 40, 1-10.
- Giot, P. (2005). Relationships Between Implied Volatility Indexes and Stock Index Returns. *Journal of Portfolio Management*, 31(3), 92-100.
- Gonzalez, L., Powell, J. G., Shi, J., & Wilson, A. (2005). Two Centuries of Bull and Bear Market Cycles. *International Review of Economics and Finance*, 14, 469-486.
- Jarque, C. M., & Bera, A. K. (1987). A Test for Normality of Observations and Regression Residuals. *International Statistical Review / Revue Internationale de Statistique*, 55(2), 163-172.
- Ljung, G. M., & Box, G. E. (1978). On a Measure of Lack of Fit in Time Series Models. *Biometrika*, 65(2), 297-303.
- Lunde, A., & Timmermann, A. (2004). Duration Dependence in Stock Prices: An Analysis of Bull and Bear Markets. *Journal of Business and Economic Statistics*, 22(3), 253-273.
- MacKinlay, A. C. (1997). Event Studies in Economics and Finance. *Journal of Economic Literature*, 35(1), 13-39.
- Odean, T. (1998). Are Investors Reluctant to Realize their Losses. *Journal of Finance*, 53(5), 1775-1798.
- Odean, T. (1999). Do Investors Trade Too Much. *American Economic Review*, 89, 1279-1298.
- Pagan, A. R., & Sossounov, K. (2003). A Simple Framework for Analysing Bull and Bear Markets. *Journal of Applied Econometrics*, 18(1), 23-46.
- S&P Dow Jones Indices . (2016). *S&P Dow Jones Indices*. Retrieved April 12, 2016, from S&P 500: http://us.spindices.com/documents/methodologies/methodology-sp-us-indices.pdf?force_download=true
- Shapiro, S. S., & Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52(3/4), 591-611.
- Shefrin, H., & Statman, M. (1985). The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence. *Journal of Finance*, 40(3), 777-790.
- Shiller, R. J. (2003). From Efficient Markets Theory to Behavioral Finance. *The Journal of Economic Perspectives*, 17(1), 83-104.
- Taeger, D., & Kuhnt, S. (2014). *Statistical Hypothesis Testing with SAS and R*. West Sussex: John Wiley & Sons.

- Tamhane, A. C., & Dunlop, D. D. (2000). *Statistics and Data Analysis: from Elementary to Intermediate*. Upper Saddle River: Prentice Hall.
- Whaley, R. E. (2000). The Investor Fear Gauge. *Journal of Portfolio Management*, 26(3), 12-17.
- Whaley, R. E. (2009). Understanding the VIX. *Journal of Portfolio Management*, 35(3), 98-105.
- Yahoo! (2016, January 25). *Yahoo!* Retrieved from Yahoo! Finance: <http://finance.yahoo.com/q?s=vix>
- Yahoo! (2016, January 25). *Yahoo! Finance*. Retrieved from Yahoo!: <http://finance.yahoo.com/q?s=gspc>
- Zakamulin, V. (2016). Abnormal Stock Market Returns around Peaks in VIX: The Evidence of Investors Overreaction? *Second draft*.

11 Appendix

11.1 R-file

11.1.1 Filename: Download-Data

```
rm(list=ls(all=TRUE))
library(tseries)
library(zoo)

start.date <- "1990-01-01"
end.date <- "2015-12-31"
filename <- "sp500vix.txt"

# Download data from the Internet

sp500 = get.hist.quote(instrument="^GSPC", start=start.date,
                      end=end.date, quote="AdjClose",
                      provider="yahoo", origin="1970-01-01",
                      compression="d", retclass="zoo")
vix = get.hist.quote(instrument="^VIX", start=start.date,
                    end=end.date, quote="AdjClose",
                    provider="yahoo", origin="1970-01-01",
                    compression="d", retclass="zoo")

data.zoo = merge(sp500,vix) # merge the data in a single zoo object
names = c("SP500","VIX")
colnames(data.zoo) <- names # assign names to columns of data

# Save to a file
write.zoo(data.zoo, file=filename)
```

11.1.2 Filename: Identifying-Algorithms

```
# BB ALGORITHMS FOR BULL-BEAR PARTITION
# =====
# Bull-Bear partition based on the Change given by some Threshold value
# LT Algorithm of Lunde and Timmermann (2004)

BBPartition.LTRules <- function(index, threshold.bull, threshold.bear) {
  # inputs:
  # index - vector of stock price index
  # threshold.bull - value of threshold used to detect a Bull state
  # threshold.bear - value of threshold used to detect a Bear state

  nobs <- length(index)
  bull <- rep(FALSE,nobs)
  enter <- vector()
  exit <- vector()
}
```

```

instocks <- FALSE

for (i in 1:nobs) {
  if (i == 1) {
    enter <- c(enter, i)
    bull[i] <- TRUE
    instocks <- TRUE
  } else {
    if (instocks) {
      # check whether we need to exit
      maxval <- max(index[enter[length(enter)]:i])
      change <- (maxval-index[i])/maxval*100
      if (change > threshold.bear) {
        exit <- c(exit, i)
        instocks <- FALSE
        # now find the BEAR period
        ind <- which.max(index[enter[length(enter)]:i]) + enter[length(enter)]
        for(j in ind:i) bull[j] <- FALSE
      } else {
        bull[i] <- TRUE
      }
    } else {
      # check whether we need to enter
      minval <- min(index[exit[length(exit)]:i])
      change <- (index[i]-minval)/minval*100
      if (change > threshold.bull) {
        instocks <- TRUE
        enter <- c(enter, i)
        # now find the BULL period
        ind <- which.min(index[exit[length(exit)]:i]) + exit[length(exit)]
        for(j in ind:i) bull[j] <- TRUE
      } else {
        bull[i] <- FALSE
      }
    }
  }
}

return(bull)
}

# =====
# Bull-Bear BB Algorithm of Bry and Boschan (1971)

# The parameters of the algorithm
roll.window.size <- 4 # the half-size of the window to find minima and maxima
margin.size <- 20 # the size of the left and right margin
min.cycle.length <- 22 # minimum full cycle length
Max.Change <- 0.3 # the change in vix/vol that invalidates the minimum cycle length
min.phase.length <- 10 # the minimum phase (bull or bear) length

```

```

getBull <- function(local.min, local.max) {
  # this is a final function that, given a sequence of local.min and local.max,
  # creates a single vector that contains TRUE for Bull periods and FALSE for BEAR periods
  # Input - vectors that contain points of local.min and local.max
  # both input vectors must be of the same length

  n <- length(local.min)
  if(length(local.min) != n) stop("Vectors that contain local min and max are of different
lengths!")
  bull <- rep(FALSE, n)

  # identify bull and bear markets
  first.ext.found <- FALSE # first extremum, either local.min or local.max, FALSE if not
found yet
  isBull <- FALSE
  for(i in 1:n) {

    # find the first local min or max
    if( (first.ext.found==FALSE) & (local.min[i]==TRUE) ) {
      # beginning is BEAR, next is BULL
      first.ext.found <- TRUE
      isBull <- TRUE
      next
    }
    if( (first.ext.found==FALSE) & (local.max[i]==TRUE) ) {
      first.ext.found <- TRUE
      # beginning is BULL, next BEAR
      # mark all previous as BULL
      for(j in 1:i) bull[j] <- TRUE
      isBull <- FALSE
      next
    }
  }

  # if the first extremum is found
  if(first.ext.found==TRUE) {
    if(isBull==TRUE) {
      bull[i] <- TRUE
      # this is bull period
      if(local.max[i]==TRUE) isBull <- FALSE # Next period BEAR
    } else {
      # this is bear period
      bull[i] <- FALSE
      if(local.min[i]==TRUE) isBull <- TRUE # Next period BULL
    }
  }
}

return(bull)
}

```

```

eliminateMultipleMM <- function(index, local.min, local.max) {

  # THIS PROCEDURE ELIMINATE MULTIPLE MAXIMA AND MINIMA IN THE
  BEGINNING, END
  # MULTIPLE MINIMA WITHIN TWO MAXIMA, MULTIPLE MAXIMA WITHIN TWO
  MINIMA

  n <- length(index)

  #=====
  # eliminating multiple minima and maxima in the beginning

  # find the first min and max
  i <- 1
  while(local.max[i]==FALSE) {
    i <- i+1
    first.max <- i
  }
  i <- 1
  while(local.min[i]==FALSE) {
    i <- i+1
    first.min <- i
  }

  # eliminate multiple minima in the beginning
  # find all minima
  minima <- vector()
  ind.minima <- vector()
  for(i in 1:first.max) {
    if(local.min[i]==TRUE) {
      minima <- c(minima,index[i])
      ind.minima <- c(ind.minima, i)
    }
  }
  # proceed with the counting the number of minima
  nMin <- length(minima)
  if(nMin >= 1) {
    if(min(minima)>index[1]) {
      # eliminate all of them
      for(i in 1:first.max) local.min[i] <- FALSE
    } else {
      # keep only the minimum of them
      ind.min <- which.min(minima)
      for(i in 1:nMin) if(i != ind.min) local.min[ind.minima[i]] <- FALSE
    }
  }

  # eliminate multiple maxima in the beginning
  # find all maxima

```

```

maxima <- vector()
ind.maxima <- vector()
for(i in 1:first.min) {
  if(local.max[i]==TRUE) {
    maxima <- c(maxima,index[i])
    ind.maxima <- c(ind.maxima, i)
  }
}
# proceed with the counting the number of maxima
nMax <- length(maxima)
if(nMax >= 1) {
  if(max(maxima)<index[1]) {
    # eliminate all of them
    for(i in 1:first.min) local.max[i] <- FALSE
  } else {
    # keep only the largest of them
    ind.max <- which.max(maxima)
    for(i in 1:nMax) if(i != ind.max) local.max[ind.maxima[i]] <- FALSE
  }
}

#=====
# eliminating multiple minima and maxima in the end

# find the last min and max
i <- n
while(local.max[i]==FALSE) {
  i <- i-1
  last.max <- i
}
i <- n
while(local.min[i]==FALSE) {
  i <- i-1
  last.min <- i
}

# find all minima
minima <- vector()
ind.minima <- vector()
for(i in last.max:n) {
  if(local.min[i]==TRUE) {
    minima <- c(minima,index[i])
    ind.minima <- c(ind.minima, i)
  }
}
# proceed with the counting the number of minima
nMin <- length(minima)
if(nMin >= 1) {
  if(min(minima)>index[n]) {
    # eliminate all of them

```

```

    for(i in last.max:n) local.min[i] <- FALSE
  } else {
    # keep only the minimum of them
    ind.min <- which.min(minima)
    for(i in 1:nMin) if(i != ind.min) local.min[ind.minima[i]] <- FALSE
  }
}

# eliminate multiple maxima in the end
# find all maxima
maxima <- vector()
ind.maxima <- vector()
for(i in last.min:n) {
  if(local.max[i]==TRUE) {
    maxima <- c(maxima,index[i])
    ind.maxima <- c(ind.maxima, i)
  }
}
# proceed with the counting the number of maxima
nMax <- length(maxima)
if(nMax >= 1) {
  if(max(maxima)<index[n]) {
    # eliminate all of them
    for(i in last.min:n) local.max[i] <- FALSE
  } else {
    # keep only the largest of them
    ind.max <- which.max(maxima)
    for(i in 1:nMax) if(i != ind.max) local.max[ind.maxima[i]] <- FALSE
  }
}
}

```

```

#=====

```

```

# SECOND PART, Eliminate multiple max within two minima and
# multiple min within two maxima

```

```

# eliminate multiple minima within two maxima
first <- FALSE
ind.first <- 1
for(i in 1:n) {
  if( (local.max[i]==TRUE) & (first==FALSE) ) {
    # this is the very first maximum
    first <- TRUE
    ind.first <- i
    next
  }
  if((local.max[i]==TRUE) & (first==TRUE)) {
    # this is the subsequent max
    # elimination of multiple minima within ind.first and i
    # start with the search for minima

```

```

minima <- vector()
ind.minima <- vector()
for(j in ind.first:i) {
  if(local.min[j]==TRUE) {
    minima <- c(minima,index[j])
    ind.minima <- c(ind.minima, j)
  }
}
# proceed with the counting the number of minima
nMin <- length(minima)
# eliminate multiple minima
if(nMin > 1) {
  ind.min <- which.min(minima)
  for(s in 1:nMin) if(s != ind.min) local.min[ind.minima[s]] <- FALSE
}
# end of procedure
ind.first <- i
}
}

# eliminate multiple maxima within two minima
first <- FALSE
ind.first <- 1
for(i in 1:n) {
  if( (local.min[i]==TRUE) & (first==FALSE) ) {
    # this is the very first maximum
    first <- TRUE
    ind.first <- i
    next
  }
  if((local.min[i]==TRUE) & (first==TRUE)) {
    # this is the subsequent max
    # elimination of multiple minima within ind.first and i
    # start with the search for minima
    maxima <- vector()
    ind.maxima <- vector()
    for(j in ind.first:i) {
      if(local.max[j]==TRUE) {
        maxima <- c(maxima,index[j])
        ind.maxima <- c(ind.maxima, j)
      }
    }
  }
  # proceed with the counting the number of minima
  nMax <- length(maxima)
  # eliminate multiple minima
  if(nMax > 1) {
    ind.max <- which.max(maxima)
    for(s in 1:nMax) if(s != ind.max) local.max[ind.maxima[s]] <- FALSE
  }
  # end of procedure
}
}

```

```

    ind.first <- i
  }
}
return(list(l.min=local.min, l.max=local.max))
}

```

```

Partition.BBRules <- function(index) {

```

```

  # MAIN function that find the Bull and Bear states
  # uses the dating rules that are described in Pagan and Sousunov
  # inputs:
  # index - the vector of stock price index

```

```

  k <- roll.window.size

```

```

  p.length <- 2*k + 1
  if(p.length > min.phase.length)
    stop("Rolling window size is greater than the minimum phase length! \n")

```

```

  n <- length(index)
  local.min <- rep(FALSE,n)
  local.max <- rep(FALSE,n)

```

```

  # first run, identify min and max in a rolling window

```

```

  for(i in (k+1):(n-k)) {
    start <- max(1, i - k)
    end <- min(n, i + k)
    # find local max
    ind <- which.max(index[start:end])

```

```

    # if the center of the window is the maximum in the window
    if(ind == (k+1)) local.max[i] <- TRUE

```

```

    #local.max[ind + start - 1] <- TRUE # old

```

```

    # find local min
    ind <- which.min(index[start:end])

```

```

    if(ind == (k+1)) local.min[i] <- TRUE

```

```

    #local.min[ind + start - 1] <- TRUE # old
  }

```

```

  # second run, elimination of successive min and max
  # need to find a sequence, choose the only one point among all
  multiple.max <- FALSE
  multiple.min <- FALSE

```

```

  first.max <- 1
  first.min <- 1

```



```

last.max <- n
last.min <- n

for(i in 2:n) {
  # markt the start of a new max (probably series)
  if( (local.max[i-1]==FALSE) & (local.max[i]==TRUE) ) first.max <- i
  # markt the existence of multiple max
  if( (local.max[i-1]==TRUE) & (local.max[i]==TRUE) ) multiple.max <- TRUE
  # check whether the end of a sequence of maxima
  if( (local.max[i]==FALSE) & (multiple.max==TRUE) ) {
    # this is the end of the sequence
    multiple.max <- FALSE
    ind <- which.max(index[first.max:(i-1)])
    # eliminate all maxima
    for(j in first.max:(i-1)) local.max[j] <- FALSE
    local.max[first.max + ind - 1] <- TRUE
  }

  # markt the start of a new min (probably series)
  if( (local.min[i-1]==FALSE) & (local.min[i]==TRUE) ) first.min <- i
  # markt the existence of multiple min
  if( (local.min[i-1]==TRUE) & (local.min[i]==TRUE) ) multiple.min <- TRUE
  # check whether the end of a sequence of minima
  if( (local.min[i]==FALSE) & (multiple.min==TRUE) ) {
    # this is the end of the sequence
    multiple.min <- FALSE
    ind <- which.min(index[first.min:(i-1)])
    # eliminate all minima
    for(j in first.min:(i-1)) local.min[j] <- FALSE
    local.min[first.min + ind - 1] <- TRUE
  }
}

# first sensoring, eliminating turns within 6 months of beginning and end
w <- margin.size
for(i in 1:w) {
  local.min[i] <- FALSE
  local.max[i] <- FALSE
}
for(i in (n-w+1):n) {
  local.min[i] <- FALSE
  local.max[i] <- FALSE
}

# second sensoring,
# eliminate multiple maxima in the beginnig, end, and within two minima
# eliminate multiple minima in the beginnig, end, and within two maxima
res <- eliminateMultipleMM(index, local.min, local.max)
local.max <- res$l.max
local.min <- res$l.min

```

```

# eliminate phases whose duration is less than 4 months (except when change > 20%)
Elimination <- TRUE
while(Elimination) {
  first.max <- FALSE
  first.min <- FALSE
  ind.max <- 1
  ind.min <- 1
  for(i in 1:n) {
    # find the first point, identify it as max or min
    if((first.max==FALSE) & (first.min==FALSE) & (local.min[i]==TRUE)) {
      first.min <- TRUE
      ind.min <- i
      next
    }
    if((first.max==FALSE) & (first.min==FALSE) & (local.max[i]== TRUE)) {
      first.max <- TRUE
      ind.max <- i
      next
    }
    if((first.max==TRUE) & (local.min[i]==TRUE)) {
      # find out if we need to eliminate the local min
      change <- abs( (index[ind.max]-index[i])/index[ind.max])
      phase.length <- i - ind.max
      if( (phase.length<min.phase.length) & (change<Max.Change) ) {
        # we need to eliminate
        #local.max[ind.max] <- FALSE
        local.min[i] <- FALSE
        # sensor the rest of the points
        res <- eliminateMultipleMM(index, local.min, local.max)
        local.max <- res$l.max
        local.min <- res$l.min
        # go out of the for loop and begin again
        break
      } else {
        first.max <- FALSE
        first.min <- TRUE
        ind.min <- i
      }
    }
  }
  if((first.min==TRUE) & (local.max[i]==TRUE)) {
    # find out if we need to eliminate the first minimum
    change <- abs( (index[ind.min]-index[i])/index[ind.min])
    phase.length <- i - ind.min
    if( (phase.length<min.phase.length) & (change<Max.Change) ) {
      # we need to eliminate both
      #local.min[ind.min] <- FALSE
      local.max[i] <- FALSE
      # sensor the rest of the points
      res <- eliminateMultipleMM(index, local.min, local.max)
    }
  }
}

```

```

    local.max <- res$l.max
    local.min <- res$l.min
    # go out of the for loop and begin again
    break
  } else {
    first.min <- FALSE
    first.max <- TRUE
    ind.max <- i
  }
}

# if we go to the very end then there were no eliminations
if(i==n) Elimination <- FALSE
}
}

# FINAL PART, ELIMINATE THE CYCLES WHOSE DURATION IS LESS THAN 16
MONHTS
# EXCEP WHEN CHANGE > 20%

Elimination <- TRUE
while(Elimination) {

  first.max <- FALSE
  first.min <- FALSE
  cycle.points <- vector() # point values
  cycle.ind <- vector() # indices of points

  for(i in 1:n) {
    # find the first point, identify it as max or min
    if((first.max==FALSE) & (first.min==FALSE) & (local.min[i]==TRUE)) {
      first.min <- TRUE
      cycle.points <- c(cycle.points, index[i])
      cycle.ind <- c(cycle.ind, i)
      next
    }
    if((first.max==FALSE) & (first.min==FALSE) & (local.max[i]== TRUE)) {
      first.max <- TRUE
      cycle.points <- c(cycle.points, index[i])
      cycle.ind <- c(cycle.ind, i)
      next
    }
  }

  if(first.max == TRUE) {
    if(local.min[i] == TRUE) {
      # add minimum point
      cycle.points <- c(cycle.points, index[i])
      cycle.ind <- c(cycle.ind, i)
      next
    }
  }
}

```

```

if(local.max[i] == TRUE) {
  cycle.points <- c(cycle.points, index[i])
  cycle.ind <- c(cycle.ind, i)
  # we need to check if we need to eliminate the first point
  cycle.length <- cycle.ind[3]-cycle.ind[1]+1
  change.1 <- abs( (cycle.points[2]-cycle.points[1])/cycle.points[1])
  change.2 <- abs( (cycle.points[3]-cycle.points[2])/cycle.points[2])
  max.change <- max(change.1, change.2)
  if( (cycle.length < min.cycle.length) & (max.change < Max.Change) ) {
    # need to eliminate
    local.max[cycle.ind[1]] <- FALSE
    # sensor the rest of the points
    res <- eliminateMultipleMM(index, local.min, local.max)
    local.max <- res$l.max
    local.min <- res$l.min
    # go out of the for loop and begin again
    break
  } else {
    # do not need to eliminate, proceed further
    first.max <- FALSE
    first.min <- TRUE
    cycle.points <- cycle.points[-1]
    cycle.ind <- cycle.ind[-1]
    next
  }
}
}
}

if(first.min == TRUE) {
  if(local.max[i] == TRUE) {
    # add maximum point
    cycle.points <- c(cycle.points, index[i])
    cycle.ind <- c(cycle.ind, i)
    next
  }
  if(local.min[i] == TRUE) {
    cycle.points <- c(cycle.points, index[i])
    cycle.ind <- c(cycle.ind, i)
    # we need to check if we need to eliminate the first point
    cycle.length <- cycle.ind[3]-cycle.ind[1]+1
    change.1 <- abs( (cycle.points[2]-cycle.points[1])/cycle.points[1])
    change.2 <- abs( (cycle.points[3]-cycle.points[2])/cycle.points[2])
    max.change <- max(change.1, change.2)
    if( (cycle.length < min.cycle.length) & (max.change < Max.Change) ) {
      # need to eliminate
      local.min[cycle.ind[1]] <- FALSE
      # sensor the rest of the points
      res <- eliminateMultipleMM(index, local.min, local.max)
      local.max <- res$l.max
      local.min <- res$l.min
    }
  }
}
}

```

```

    # go out of the for loop and begin again
    break
  } else {
    # do not need to eliminate, proceed further
    first.max <- TRUE
    first.min <- FALSE
    cycle.points <- cycle.points[-1]
    cycle.ind <- cycle.ind[-1]
    next
  }
}
}

# if we go to the very end then there were no eliminations
if(i==n) Elimination <- FALSE
} # end for-loop
}

bull <- getBull(local.min, local.max)
return(bull)
}

```

11.1.3 Filename: Data

```

rm(list=ls(all=TRUE))
library(zoo)
library(ggplot2)
source("Identifying-Algorithms.r")

process.voldata <- function(vol, dates, method=c("BB", "LT")) {
  # this function determines the bull and bear periods in vol
  # and returns the starts and ends of Bear periods

  # inputs
  # - vol - vector containing the variable used to determine bull and bear states
  # - dates - the vector of dates
  # - method - defines the method of BB of bull and bear states

  switch(method,
    "BB" = {
      # BB Algorithm of Bry and Boschan (1971) #WE ARE NOT USING SMOOTHED
DATA
      bull <- Partition.BBRules(vol)
    },
    "LT" = {
      # LT Algorithm of Lunde and Timmermann (2004)
      threshold.bull <- 30
      threshold.bear <- 20
      bull <- BBPartition.LTRules(vol, threshold.bull, threshold.bear)
    },
  )
}

```

```

    {
      stop("Unknown method!")
    }
  )

# create the rectangles for the BEAR shaded areas
xstart <- vector()
xend <- vector()
n <- length(bull)
isBear <- FALSE
for(i in 1:n) {
  if((bull[i] == FALSE) & (isBear == FALSE)) {
    xstart <- c(xstart, dates[i])
    isBear <- TRUE
  }
  if((bull[i] == TRUE) & (isBear == TRUE)) {
    xend <- c(xend, dates[i])
    isBear <- FALSE
  }
  if((i == n) & (isBear == TRUE)) xend <- c(xend, dates[i])
}

# creates rectangle areas
rects <- data.frame(xstart=as.Date(xstart), xend=as.Date(xend))
return(list(vol=vol, bull=bull, rects=rects))
}

divide.ret <- function(ret, bull) {
# this function separates the returns into
# (1) - returns during Increasing vol and (2) returns during decreasing vol

# inputs
# ret - vector of returns
# bull - vector of TRUE (for bull state) and FALSE (for bear states)

volup <- vector()
voldn <- vector()
n <- length(bull)
for(i in 1:n) {
  if(bull[i]==TRUE) {
    volup <- c(volup, ret[i])
  } else {
    voldn <- c(voldn, ret[i])
  }
}
return(list(volup=volup, voldn=voldn))
}

# Read the data
filename <- "sp500vix.txt"

```

```

data.zoo <- read.zoo(file=filename, header = TRUE)

# Select the required historical period
Period <- "Full"
#Period <- "First"
#Period <- "Second"

switch(Period,
      "Full" = {
        start.date <- "1990-01-01"
        end.date <- "2015-12-31"
        data.zoo <- window(data.zoo, start=start.date, end=end.date)
      },
      "First" = {
        start.date <- "1990-01-01"
        end.date <- "2002-12-31"
        data.zoo <- window(data.zoo, start=start.date, end=end.date)
      },
      "Second" = {
        start.date <- "2003-01-01"
        end.date <- "2015-12-31"
        data.zoo <- window(data.zoo, start=start.date, end=end.date)
      },
      {
        stop("Unknown period!")
      }
)

dates <- index(data.zoo) # retrieve dates from zoo-object
data <- coredata(data.zoo) # retrieve price data from zoo-object
nobs <- nrow(data)

# Compute returns
prices <- data[,"SP500"]
#ret <- diff(prices)/prices[1:(nobs-1)] # STANDARD RETURNS
ret <- diff(log(prices)) # LOG RETURNS

vix <- data[,"VIX"]
rvix <- diff(vix)/vix[1:nobs-1]

# skip the first observation
vix <- data[2:nobs,"VIX"]
dates <- dates[2:nobs]
prices <- prices[2:nobs]

# construct time-series objects and plot the data
names = c("SP500","VIX")
data <- cbind(log(prices),vix)
data.zoo <- zoo(data, order.by=dates)
colnames(data.zoo) <- names # assign names to columns of data

```

```

plot(data.zoo)

# Find the periods of Increasing and Decreasing Vol
res.vix <- process.voldata(vix, dates, method="BB")
#res.vix <- process.voldata(vix, dates, method="LT")

vix <- res.vix$vol
bull <- res.vix$bull
rects <- res.vix$rects

# plot the index
df <- data.frame(Date=as.Date(dates), Value=vix)

ggplot() +
  geom_rect(data = rects, aes(xmin = xstart, xmax = xend, ymin = -Inf, ymax = Inf),
  fill="gray") +
  geom_line(data = df, aes(Date,Value)) +
  theme_bw() + xlab("") + ylab("VIX")

# test of equality of mean returns during Increasing and Decreasing Vol
res.vix <- divide.ret(ret, bull)
volup <- res.vix$volup
voldn <- res.vix$voldn
t.test(volup,voldn)
wilcox.test(volup,voldn)

```

11.1.4 Filename: Algorithm-BB

```

rm(list=ls(all=TRUE))
library(zoo)
library(tseries)
library(ggplot2)
library(reshape2)
library(plyr)
library(xts)
library(data.table)
source("Data.r")

### SUMMARY STATISTICS of S&P 500 log-returns ###
min(ret)*100 #minimum
mean(ret)*100 #mean
t.test(ret) #p-value H0:my=0
sd(ret)*100 #std. deviation
max(ret)*100 #maximum
skewness(ret) #skewness
kurtosis(ret) #kurtosis
#shapiro.test(ret) #Shapiro-Wilk normality test, p-value
acf(ret,plot=F,lag.max=1) #autocorrelation, one lag
Box.test(ret,lag=1,type = "Ljung") #p-value H0:lag1=0

```



```

#### FIRST STUDY ####
#Regression
fitsym2000 <- lm(ret ~ rvix) #symmetric
summary(fitsym2000)

fitsym2009 <- lm(rvix ~ ret) #symmetric
summary(fitsym2009)

rvixplus <- vector()
for (i in 1:length(rvix)){
  if(rvix[i]>0){
    rvixplus[i] <- rvix[i]
  }else{
    rvixplus[i] <- 0
  }
}

retminus <- vector()
for (i in 1:length(ret)){
  if(ret[i]<0){
    retminus[i] <- ret[i]
  }else{
    retminus[i] <- 0
  }
}

asymfit2000 <- lm(ret ~ rvix + rvixplus)
summary(asymfit2000) #asymmetric

asymfit2009 <- lm(rvix ~ ret + retminus)
summary(asymfit2009) #asymmetric

#### SECOND STUDY ####
#Finding peaks with VIX
ipeak <- vector()
for (i in 2:length(bull)){
  if ((bull[i]==FALSE) & (bull[i-1]==TRUE))
    ipeak <- c(ipeak,i-1)
}
print(ipeak)

# Finding valleys with VIX
ivalley <- vector()
for(i in 2:length(bull)){
  if ((bull[i]==TRUE) & (bull[i-1]==FALSE))
    ivalley <- c(ivalley,i-1)
}

```

```

print(ivalley)

#### STATISTICS, RISING/FALLING for VIX ####
# TRUE = 1 and FALSE = 0
bull01 <- bull*1
#bull01

# DURATION #
durdf <- data.frame(Date=(dates), Value1=bull01)

setDT(durdf)
durdf[, run := cumsum(c(1, diff(Value1) !=0))]

unique(durdf$Value1)
one <- subset(durdf, Value1=="1")
zero <- subset(durdf, Value1=="0")
nrow(one) #Total days of rising
nrow(zero) #Total days of falling

#Total period
duration <- rep(0)
for (i in 1:295){
  ind <- which(durdf$run==i)
  a <- durdf$Date[ind]
  duration[i] <- length(a)
}

c <- rep(c(1,0),295)
c <- c[1:295]
durdf2 <- data.frame(duration, type=c)

# #First period
# duration <- rep(0)
# for (i in 1:143){
#   ind <- which(df$run==i)
#   a <- df$Date[ind]
#   duration[i] <- length(a)
# }
# c <- rep(c(1,0),143)
# c <- c[1:143]
# df2 <- data.frame(duration, type=c)

#Second period
# duration <- rep(0)
# for (i in 1:151){
#   ind <- which(df$run==i)
#   a <- df$Date[ind]
#   duration[i] <- length(a)
# }

```

```

# c <- rep(c(1,0),151)
# c <- c[1:151]
# df2 <- data.frame(duration, type=c)

unique(durdf2$type)
r <- subset(durdf2, c=="1")
f <- subset(durdf2, c=="0")

#rising period
nrow(r) #phases
min(r$duration)
mean(r$duration)
median(r$duration)
max(r$duration)

#falling period
nrow(f) #phases
min(f$duration)
mean(f$duration)
median(f$duration)
max(f$duration)

# AMPLITUDE #
mean(vix)
print(vix[1], digits=6)
vixpeak <- vix[ipeak]
vixvalley <- vix[ivalley]

#Falling amplitude
amplitudedef <- vector(mode="numeric", length=length(vixvalley))

for(i in 1:length(vixpeak)){
  amplitudedef[i] <- abs((vixvalley[i]-vixpeak[i])/vixpeak[i])
}

list(amplitudedef, sort(amplitudedef))
print(length(amplitudedef))
min(amplitudedef)
mean(amplitudedef)
median(amplitudedef)
max(amplitudedef)

print(median(amplitudedef))

#Raising amplitude
amplituder <- vector(mode="numeric", length=length(vixvalley)-1)

i <- 0
while(i < length(vixvalley) && i < length(vixpeak)+1 )
{

```

```

    amplituder[i] <- abs((vixpeak[i+1]-vixvalley[i])/vixvalley[i])
    i <- i+1
  }

print(length(amplituder))
min(amplituder)
mean(amplituder)
median(amplituder)
max(amplituder)

# RISING statistics #
cumr <- cumsum(volup) #average cum. return
last(cumr)
mean(volup)*100#mean daily return
t.test(volup) #p-value H0:my=0
sd(volup)*100 #daily std. deviation
skewness(volup) #skewness
kurtosis(volup) #kurtosis
shapiro.test(volup) #Shapiro-Wilk normality test, p-value
acf(volup,plot=F,lag.max=1) #autocorrelation, one lag
Box.test(volup,lag=1, type = "Ljung") #p-value H0:lag1=0

# FALLING statistics #
cumf <- cumsum(voldn)#average cum. return
last(cumf)
mean(voldn)*100 #mean daily return
t.test(voldn) #p-value H0:my=0
sd(voldn)*100 #daily std. deviation
skewness(voldn) #skewness
kurtosis(voldn) #kurtosis
shapiro.test(voldn) #Shapiro-Wilk test, p-value
acf(voldn,plot=F,lag.max=1) #autocorrelation, one lag
Box.test(voldn,lag=1, type = "Ljung") #p-value H0:lag1=0

### THIRD STUDY ###
#Abnormal returns
mea <- vector()
n <- length(bull)
for(i in 1:n) {
  if(bull[i]==TRUE) {
    mea <- c(mea, mean(ret))
  } else {
    mea <- c(mea, mean(ret))
  }
}
abret <- vector()
n <- length(bull)
for(i in 1:n) {
  if(bull[i]==TRUE) {

```

```

    abret[i] <- ret[i]-mea[i]
  } else {
    abret[i] <- ret[i]-mea[i]
  }
}

u <- (abret+mea)

v <- vector()
for (i in 1:length(bull)){
  if(u[i]==ret[i]){
    v[i] <- 1
  }else{
    v[i] <- 0
  }
}

compare <- data.frame(abret,mea,u,ret,v,bull01)

# FINDING RETURNS of n-days before and after a peak #
n <- 5 #using 5, 10, 15, and 20 days
before <- vector()
after <- vector()
for (i in 1:length(ipeak)){
  before <- c(before,abret[(ipeak[i]-n):(ipeak[i]-1)])
  after <- c(after,abret[(ipeak[i]+1):(ipeak[i]+n)])
}

mean(before)*100
mean(after,na.rm=TRUE)*100

#tests
t.test(before,after)
wilcox.test(before,after)
ks.test(before,after)

#Event study using abnormal returns
source("window.table.BB.R")

# FIGURES #
# ABNORMAL RETURN AND CUMULATIVE ABNORMAL RETURN #
# #Abnormal returns over the event window - total period.
# x <- -10:10 #days relative to a volatility peak
# AR <- read.table("AR.total.txt",sep="\t", col.names=c("AR"), fill=FALSE,
strip.white=TRUE)
# t.test(AR)
#
# df95 <- data.frame(x, AR)
# df95$lb <- df95$AR-0.2675931
# df95$sub <- df95$AR+0.1737227

```

```

#
# ggplot(data = df95, aes(x = x, y = AR)) +
#   ylab("Abnormal Return, %") + xlab("Days relative to a volatility peak") +
#   geom_line(size = 0.5) +
#   geom_hline(aes(yintercept = 0), linetype = "dashed") +
#   geom_ribbon(aes(ymin=lb, ymax=ub), alpha=0.5) +
#   scale_x_continuous(breaks = c(-10,-9,-8,-7,-6,-5,-4,-3,-2,-1,0,1,2,3,4,5,6,7,8,9,10)) +
#   scale_y_continuous(breaks = c(-1,0,1))
#
# #Cumulative Abnormal Returns over the event window - total period.
# x <- -10:10 #days relative to a volatility peak
# CAR <- read.table("CAR.total.txt",sep="\t", col.names=c("CAR"), fill=FALSE,
strip.white=TRUE)
# t.test(CAR)
#
# df95.CAR <- data.frame(x, CAR)
# df95.CAR$lb <- df95.CAR$CAR-(-1.983744+1.520496)
# df95.CAR$ub <- df95.CAR$CAR-(-1.057248+1.520496)
#
# ggplot(data = df95.CAR, aes(x = x, y = CAR)) +
#   ylab("Cumulative Abnormal Return, %") + xlab("Days relative to a volatility peak") +
#   geom_line(size = 0.5) +
#   geom_hline(aes(yintercept = 0), linetype = "dashed") +
#   geom_ribbon(aes(ymin=lb, ymax=ub), alpha=0.5) +
#   scale_x_continuous(breaks = c(-10,-9,-8,-7,-6,-5,-4,-3,-2,-1,0,1,2,3,4,5,6,7,8,9,10)) +
#   scale_y_continuous(breaks = c(-1,0,1))
#
# #Abnormal Return over the event window. Total-, 1st- and 2nd period.
# Total.period = read.table("AR.total.txt",sep="\t", col.names=c("Total period"), fill=FALSE,
strip.white=TRUE)
# First.period = read.table("AR.first.txt",sep="\t", col.names=c("First period"), fill=FALSE,
strip.white=TRUE)
# Second.period = read.table("AR.second.txt",sep="\t", col.names=c("Second period"),
fill=FALSE, strip.white=TRUE)
#
# x <- -10:10 #days relative to a volatility peak
# y1 <- Total.period #total period
# y2 <- First.period #1st subperiod
# y3 <- Second.period #2nd subperiod
# df <- data.frame(x,y1,y2,y3)
#
# df2 <- melt(data = df, variable.name = "Period", id.vars = "x")
#
# ggplot(data = df2, aes(x = x, y = value, group = Period, color = Period)) +
#   ylab("Abnormal Return, %") + xlab("Days relative to a volatility peak") +
#   geom_line(size = 0.5) +
#   geom_point(size = 2, shape = 21, fill = "white") +
#   geom_hline(aes(yintercept = 0), linetype = "dashed") +
#   scale_x_continuous(breaks = c(-10,-9,-8,-7,-6,-5,-4,-3,-2,-1,0,1,2,3,4,5,6,7,8,9,10)) +
#   scale_y_continuous(breaks = c(-1,0,1)) +

```

```

# scale_color_discrete(labels = c("Total", "First", "Second")) +
# theme(legend.position="bottom")
#
# #Cumulative Abnormal Return over the event window. Total-, 1st- and 2nd period.
# Total.period = read.table("CAR.total.txt",sep="\t", col.names=c("Total period"),
fill=FALSE, strip.white=TRUE)
# First.period = read.table("CAR.first.txt",sep="\t", col.names=c("First period"), fill=FALSE,
strip.white=TRUE)
# Second.period = read.table("CAR.second.txt",sep="\t", col.names=c("Second period"),
fill=FALSE, strip.white=TRUE)
#
# x <- -10:10 #days relative to a volatility peak
# y4 <- Total.period #total period
# y5 <- First.period #1st subperiod
# y6 <- Second.period #2nd subperiod
# df1 <- data.frame(x,y4,y5,y6)
#
# df3 <- melt(data = df1, variable.name = "Period", id.vars = "x")
#
# ggplot(data = df3, aes(x = x, y = value, group = Period, color = Period)) +
# ylab("Cumulative Abnormal Return, %") + xlab("Days relative to a volatility peak") +
# geom_line() +
# geom_point(size = 2, shape = 21, fill = "white") +
# geom_hline(aes(yintercept = 0), linetype = "dashed") +
# scale_x_continuous(breaks = c(-10,-9,-8,-7,-6,-5,-4,-3,-2,-1,0,1,2,3,4,5,6,7,8,9,10)) +
# scale_y_continuous(breaks = c(-4,-3,-2,-1,0)) +
# scale_color_discrete(labels = c("Total", "First", "Second")) +
# theme(legend.position="bottom")
#
# #AR - total period. Above and below median of amplitude
# x <- -10:10 #days relative to a volatility peak
# AR <- read.table("AR.total.txt",sep="\t", col.names=c("AR"), fill=FALSE,
strip.white=TRUE)
# AR.above <- read.table("AR.above.txt",sep="\t", col.names=c("Above"), fill=FALSE,
strip.white=TRUE)
# AR.below <- read.table("AR.below.txt",sep="\t", col.names=c("Below"), fill=FALSE,
strip.white=TRUE)
#
# df4 <- data.frame(x,AR,AR.above,AR.below)
# df5 <- melt(data=df4, variable.name = "Amplitude", id.vars="x")
#
# ggplot(data = df5, aes(x = x, y = value, group = Amplitude, color = Amplitude)) +
# ylab("Abnormal Return, %") + xlab("Days relative to a volatility peak") +
# geom_line() +
# geom_point(size = 2, shape = 21, fill = "white") +
# geom_hline(aes(yintercept = 0), linetype = "dashed") +
# scale_x_continuous(breaks = c(-10,-9,-8,-7,-6,-5,-4,-3,-2,-1,0,1,2,3,4,5,6,7,8,9,10)) +
# scale_y_continuous(breaks = c(-6,-5,-4,-3,-2,-1,0,1,2)) +
# scale_color_discrete(labels = c("All amplitudes", "Above median", "Below median")) +
# theme(legend.position="bottom")

```

```

#Divide amplitude in above and below median of raising amplitude
ipeak2 <- vector()
for (i in 1:length(ipeak)){
  ipeak2[i] <- ipeak[i]
}
ipeak2 <- ipeak2[-1]
DFpeakamp <- data.frame(ipeak2,amplituder)
abind <- (rep(c(0,1),73))
abind <- c[1:146]
abind <- sort(abind) #order 0 and 1 as an indicator for below and above median amp.
new <- data.frame(DFpeakamp[ order(DFpeakamp$amplituder), ], abind)

ampbelow <- subset(new$ipeak2,abind=="0")
ampabove <- subset(new$ipeak2,abind=="1")

# FUGURES #
# ABOVE MEDIAN AND BELOW MEDIAN #
# #CAR - total period. Above and below median of amplitude
# x <- -10:10 #days relative to a volatility peak
# CAR <- read.table("CAR.total.txt",sep="\t", col.names=c("CAR"), fill=FALSE,
strip.white=TRUE)
# CAR.above <- read.table("CAR.above.txt",sep="\t", col.names=c("Above"), fill=FALSE,
strip.white=TRUE)
# CAR.below <- read.table("CAR.below.txt",sep="\t", col.names=c("Below"), fill=FALSE,
strip.white=TRUE)
#
# df6 <- data.frame(x,CAR,CAR.above,CAR.below)
# df7 <- melt(data=df6, variable.name = "Amplitude", id.vars="x")
#
# ggplot(data = df7, aes(x = x, y = value, group = Amplitude, color = Amplitude)) +
#   ylab("Cumulative Abnormal Return, %") + xlab("Days relative to a volatility peak") +
#   geom_line() +
#   geom_point(size = 2, shape = 21, fill = "white") +
#   geom_hline(aes(yintercept = 0), linetype = "dashed") +
#   scale_x_continuous(breaks = c(-10,-9,-8,-7,-6,-5,-4,-3,-2,-1,0,1,2,3,4,5,6,7,8,9,10)) +
#   scale_y_continuous(breaks = c(-6,-5,-4,-3,-2,-1,0,1,2)) +
#   scale_color_discrete(labels = c("All amplitudes", "Above median", "Below median")) +
#   theme(legend.position="bottom")

```

11.1.5 Filname: Window.Table.Dummy.BB

```

### THIRD STUDY ###
source("Algorithm-BB.R")
# Algorithm by Bry and Boschan (BB) #
#full,first and second period we use length=ipeak
#above and below median we use length=ampabove and length=ampbelow

window.tableR <- matrix(NA,nrow = length(ipeak),ncol = 11)

```



```

colnames(window.tableR) <- c("-10":"0")

window.tableR[,1] <- (ret-mean(ret))[ipeak-10]
window.tableR[,2] <- (ret-mean(ret))[ipeak-9]
window.tableR[,3] <- (ret-mean(ret))[ipeak-8]
window.tableR[,4] <- (ret-mean(ret))[ipeak-7]
window.tableR[,5] <- (ret-mean(ret))[ipeak-6]
window.tableR[,6] <- (ret-mean(ret))[ipeak-5]
window.tableR[,7] <- (ret-mean(ret))[ipeak-4]
window.tableR[,8] <- (ret-mean(ret))[ipeak-3]
window.tableR[,9] <- (ret-mean(ret))[ipeak-2]
window.tableR[,10] <- (ret-mean(ret))[ipeak-1]
window.tableR[,11] <- (ret-mean(ret))[ipeak]

window.tableF <- matrix(NA,nrow = length(ipeak),ncol = 10)
colnames(window.tableF) <- c("1":"10")

window.tableF[,1] <- (ret-mean(ret))[ipeak+1]
window.tableF[,2] <- (ret-mean(ret))[ipeak+2]
window.tableF[,3] <- (ret-mean(ret))[ipeak+3]
window.tableF[,4] <- (ret-mean(ret))[ipeak+4]
window.tableF[,5] <- (ret-mean(ret))[ipeak+5]
window.tableF[,6] <- (ret-mean(ret))[ipeak+6]
window.tableF[,7] <- (ret-mean(ret))[ipeak+7]
window.tableF[,8] <- (ret-mean(ret))[ipeak+8]
window.tableF[,9] <- (ret-mean(ret))[ipeak+9]
window.tableF[,10] <- (ret-mean(ret))[ipeak+10]

#window.bull <- c(bull[ipeak-10])

window.bull <- matrix(NA,nrow = length(ipeak),ncol = 11)
colnames(window.bull)<-c("-10":"0")
window.bull[,1] <- bull[ipeak-10]
window.bull[,2] <- bull[ipeak-9]
window.bull[,3] <- bull[ipeak-8]
window.bull[,4] <- bull[ipeak-7]
window.bull[,5] <- bull[ipeak-6]
window.bull[,6] <- bull[ipeak-5]
window.bull[,7] <- bull[ipeak-4]
window.bull[,8] <- bull[ipeak-3]
window.bull[,9] <- bull[ipeak-2]
window.bull[,10] <- bull[ipeak-1]
window.bull[,11] <- bull[ipeak]

window.bear <- matrix(NA,nrow = length(ipeak),ncol = 10)
colnames(window.bear)<-c("1":"10")

window.bear[,1] <- bull[ipeak+1]
window.bear[,2] <- bull[ipeak+2]
window.bear[,3] <- bull[ipeak+3]

```

```

window.bear[,4] <- bull[ipeak+4]
window.bear[,5] <- bull[ipeak+5]
window.bear[,6] <- bull[ipeak+6]
window.bear[,7] <- bull[ipeak+7]
window.bear[,8] <- bull[ipeak+8]
window.bear[,9] <- bull[ipeak+9]
window.bear[,10] <- bull[ipeak+10]

```

```

for (i in 1:length(ipeak)){
  for(j in 1:11){
    if(window.bull[i,j]==FALSE){
      window.tableR[i,j] <- NA
    }
  }
}

```

```

# #BB full period

```

```

# window.tableR[70,1] <- NA #fix TURE values that are not close to the peak (FALSE
inbetween the TRUE and peak)
# window.tableR[70,2] <- NA
# window.tableR[111,1] <- NA
# window.tableR[111,2] <- NA
# window.tableR[119,1] <- NA
# window.tableR[119,2] <- NA
# window.tableR[142,1] <- NA

```

```

# #BB first period

```

```

# window.tableR[70,1] <- NA
# window.tableR[70,2] <- NA

```

```

# #BB second period

```

```

# window.tableR[39,1] <- NA
# window.tableR[39,2] <- NA
# window.tableR[47,1] <- NA
# window.tableR[47,2] <- NA
# window.tableR[70,1] <- NA

```

```

# #BB full period above median

```

```

# window.tableR[11,1] <- NA
# window.tableR[11,2] <- NA
# window.tableR[54,1] <- NA
# window.tableR[54,2] <- NA

```

```

# #BB full period below median

```

```

# window.tableR[60,1] <- NA
# window.tableR[61,1] <- NA
# window.tableR[61,2] <- NA

```

```

for (i in 1:length(ipeak)){
  for(j in 1:10){
    if(window.bear[i,j]==TRUE){
      window.tableF[i,j] <- NA
    }
  }
}

# #BB full period
# window.tableF[69,10] <- NA #Fix FALSE values that are not close to the (TRUE
inbetween the FALSE and peak)
# window.tableF[110,10] <- NA
# window.tableF[118,10] <- NA

# #BB first period
# window.tableF[69,10] <- NA

# #BB second period
# window.tableF[38,10] <- NA
# window.tableF[46,10] <- NA

# #BB full period above median
# window.tableF[68,10] <- NA
# window.tableF[69,10] <- NA
# window.tableF[71,10] <- NA

window.table <- matrix(NA,nrow = length(ipeak),ncol = 21)
colnames(window.table) <- c("-10":"10")

window.table[,1:11] <- window.tableR
window.table[,12:21] <- window.tableF

AR.total <- apply(window.table,2,mean,na.rm=TRUE)*100 #abnormal return of all peaks for
n-days before and after
AR.total
t.test(AR.total)
CAR.total <- cumsum(AR.total) #cumulative abnormal return
CAR.total
t.test(CAR.total)

#save to a file
# write.zoo(AR.total, file="AR.total.txt") #total = full period
# write.zoo(CAR.total, file="CAR.total.txt")
# write.zoo(AR.first, file="AR.first.txt") #first = 1st half
# write.zoo(CAR.first, file="CAR.first.txt")
# write.zoo(AR.second, file="AR.second.txt") #second = 2nd half
# write.zoo(CAR.second, file="CAR.second.txt")
# write.zoo(AR.above, file="AR.above.txt") #above median
# write.zoo(CAR.above, file="CAR.above.txt")

```

```
# write.zoo(AR.below, file="AR.below.txt") #below median
# write.zoo(CAR.below, file="CAR.below.txt")
```

11.1.6 Filename: Algorithm-LT

```
rm(list=ls(all=TRUE))
library(zoo)
library(tseries)
library(ggplot2)
library(reshape2)
library(plyr)
library(xts)
library(data.table)
source("Data.r")

#### SUMMARY STATISTICS ####
min(ret)*100 #minimum
mean(ret)*100 #mean
t.test(ret) #p-value H0:my=0
sd(ret)*100 #std. deviation
max(ret)*100 #maximum
skewness(ret) #skewness
kurtosis(ret) #kurtosis
#shapiro.test(ret) #Shapiro-Wilk normality test, p-value
acf(ret,plot=F,lag.max=1) #autocorrelation, one lag
Box.test(ret,lag=1, type="Ljung") #p-value H0:lag1=0

#### SECOND STUDY ####
#Finding peaks with VIX
ipeak <- vector()
for (i in 2:length(bull)){
  if ((bull[i]==FALSE) & (bull[i-1]==TRUE))
    ipeak <- c(ipeak,i-1)
}
print(ipeak)

# Finding valleys with VIX
ivalley <- vector()
for(i in 2:length(bull)){
  if ((bull[i]==TRUE) & (bull[i-1]==FALSE))
    ivalley <- c(ivalley,i-1)
}
print(ivalley)

#### STATISTICS, RISING/FALLING ####
# TRUE = 1 and FALSE = 0
bull01 <- bull*1
#bull01
```

```

# DURATION #
durdf <- data.frame(Date=(dates), Value1=bull01)

setDT(durdf)
durdf[, run := cumsum(c(1, diff(Value1) !=0))]

unique(durdf$Value1)
one <- subset(durdf, Value1=="1")
zero <- subset(durdf, Value1=="0")
nrow(one) #Total days of rising
nrow(zero) #Total days of falling

#Total period
duration <- rep(0)
for (i in 1:253){
  ind <- which(durdf$run==i)
  a <- durdf$Date[ind]
  duration[i] <- length(a)
}

c <- rep(c(1,0),253)
c <- c[1:253]
durdf2 <- data.frame(duration, type=c)

# #First period
# duration <- rep(0)
# for (i in 1:115){
#   ind <- which(df$run==i)
#   a <- df$Date[ind]
#   duration[i] <- length(a)
# }
# c <- rep(c(1,0),115)
# c <- c[1:115]
# df2 <- data.frame(duration, type=c)

#Second period
# duration <- rep(0)
# for (i in 1:138){
#   ind <- which(df$run==i)
#   a <- df$Date[ind]
#   duration[i] <- length(a)
# }
# c <- rep(c(1,0),138)
# c <- c[1:138]
# df2 <- data.frame(duration, type=c)

unique(durdf2$c)
r <- subset(durdf2, c=="1")
f <- subset(durdf2, c=="0")

```

```

#rising period
nrow(r) #phases
min(r$duration)
mean(r$duration)
median(r$duration)
max(r$duration)

#falling period
nrow(f) #phases
min(f$duration)
mean(f$duration)
median(f$duration)
max(f$duration)

# AMPLITUDE #
mean(vix)
print(vix[1], digits=6)
vixpeak <- vix[ipeak]
vixvalley <- vix[ivalley]

#Falling amplitude
amplitudedef <- vector(mode="numeric", length=length(vixvalley))

for(i in 1:length(vixpeak)){
  amplitudedef[i] <- abs((vixvalley[i]-vixpeak[i])/vixpeak[i])
}

list(amplitudedef, sort(amplitudedef))
print(length(amplitudedef))
min(amplitudedef)
mean(amplitudedef)
median(amplitudedef)
max(amplitudedef)

print(median(amplitudedef))

#Raising amplitude
amplituder <- vector(mode="numeric", length=length(vixpeak)-1)

i <- 0
while(i < length(vixvalley)+1 && i < length(vixpeak) )
{
  amplituder[i] <- abs((vixpeak[i+1]-vixvalley[i])/vixvalley[i])
  i <- i+1
}

print(length(amplituder))
min(amplituder)
mean(amplituder)
median(amplituder)

```

```

max(amplituder)

# RISING statistics #
cumr <- cumsum(volup) #average cum. return IKKE HELT RIKTIG
last(cumr)
mean(volup)*100#mean daily return
t.test(volup) #p-value H0:my=0
sd(volup)*100 #daily std. deviation
skewness(volup) #skewness
kurtosis(volup) #kurtosis
shapiro.test(volup) #Shapiro-Wilk normality test, p-value
acf(volup,plot=F,lag.max=1) #autocorrelation, one lag
Box.test(volup,lag=1, type="Ljung") #p-value H0:lag1=0

# FALLING statistics #
cumf <- cumsum(voldn)#average cum. return IKKE HELT RIKTIG
last(cumf)
mean(voldn)*100 #mean daily return
t.test(voldn) #p-value H0:my=0
sd(voldn)*100 #daily std. deviation
skewness(voldn) #skewness
kurtosis(voldn) #kurtosis
shapiro.test(voldn) #Shapiro-Wilk test, p-value
acf(voldn,plot=F,lag.max=1) #autocorrelation, one lag
Box.test(voldn,lag=1, type="Ljung") #p-value H0:lag1=0

#### THIRD STUDY ####
#Abnormal returns
mea <- vector()
n <- length(bull)
for(i in 1:n) {
  if(bull[i]==TRUE) {
    mea <- c(mea, mean(ret))
  } else {
    mea <- c(mea, mean(ret))
  }
}

abret <- vector()
n <- length(bull)
for(i in 1:n) {
  if(bull[i]==TRUE) {
    abret[i] <- ret[i]-mea[i]
  } else {
    abret[i] <- ret[i]-mea[i]
  }
}

```

```

u <- (abret+mea)

v <- vector()
for (i in 1:length(bull)){
  if(u[i]==ret[i]){
    v[i] <- 1
  }else{
    v[i] <- 0
  }
}

compare <- data.frame(abret,mea,u,ret,v,bull01)

# FINDING RETURNS of n-days before and after a peak
abret <- ret-mean(ret)
n <- 5 #using 5, 10, 15, 20
before <- vector()
after <- vector()
for (i in 1:length(ipeak)){
  before <- c(before,abret[(ipeak[i]-n):(ipeak[i]-1)])
  after <- c(after,abret[(ipeak[i]+1):(ipeak[i]+n)])
}

mean(before)*100
mean(after,na.rm=TRUE)*100

#tests
t.test(before,after)
wilcox.test(before,after)
ks.test(before,after)

#Event study using abnormal returns
source("window.table.LT.R")

# FIGURES #
# ABNORMAL RETURN AND CUMULATIVE ABNORMAL RETURN #
# #Abnormal returns over the event window - total period.
# x <- -10:10 #days relative to a volatility peak
# AR <- read.table("AR.total.LT.txt",sep="\t", col.names=c("AR"), fill=FALSE,
strip.white=TRUE)
# t.test(AR)
#
# df95 <- data.frame(x, AR)
# df95$lb <- df95$AR-0.2675931
# df95$sub <- df95$AR+0.1737227
#
# ggplot(data = df95, aes(x = x, y = AR)) +
#   ylab("Abnormal Return, %") + xlab("Days relative to a volatility peak") +
#   geom_line(size = 0.5) +
#   geom_hline(aes(yintercept = 0), linetype = "dashed") +

```



```

# geom_ribbon(aes(ymin=lb, ymax=ub), alpha=0.5) +
# scale_x_continuous(breaks = c(-10,-9,-8,-7,-6,-5,-4,-3,-2,-1,0,1,2,3,4,5,6,7,8,9,10)) +
# scale_y_continuous(breaks = c(-1,0,1))
#
# #Cumulative Abnormal Returns over the event window - total period.
# x <- -10:10 #days relative to a volatility peak
# CAR <- read.table("CAR.total.LT.txt",sep="\t", col.names=c("CAR"), fill=FALSE,
strip.white=TRUE)
# t.test(CAR)
#
# df95.CAR <- data.frame(x, CAR)
# df95.CAR$lb <- df95.CAR$CAR-(-1.983744+1.520496)
# df95.CAR$ub <- df95.CAR$CAR-(-1.057248+1.520496)
#
# ggplot(data = df95.CAR, aes(x = x, y = CAR)) +
# ylab("Cumulative Abnormal Return, %") + xlab("Days relative to a volatility peak") +
# geom_line(size = 0.5) +
# geom_hline(aes(yintercept = 0), linetype = "dashed") +
# geom_ribbon(aes(ymin=lb, ymax=ub), alpha=0.5) +
# scale_x_continuous(breaks = c(-10,-9,-8,-7,-6,-5,-4,-3,-2,-1,0,1,2,3,4,5,6,7,8,9,10)) +
# scale_y_continuous(breaks = c(-1,0,1))
#
# #Abnormal Return over the event window. Total-, 1st- and 2nd period.
# Total.period = read.table("AR.total.LT.txt",sep="\t", col.names=c("Total period"),
fill=FALSE, strip.white=TRUE)
# First.period = read.table("AR.first.LT.txt",sep="\t", col.names=c("First period"),
fill=FALSE, strip.white=TRUE)
# Second.period = read.table("AR.second.LT.txt",sep="\t", col.names=c("Second period"),
fill=FALSE, strip.white=TRUE)
#
# x <- -10:10 #days relative to a volatility peak
# y1 <- Total.period #total period
# y2 <- First.period #1st subperiod
# y3 <- Second.period #2nd subperiod
# df <- data.frame(x,y1,y2,y3)
#
# df2 <- melt(data = df, variable.name = "Period", id.vars = "x")
#
# ggplot(data = df2, aes(x = x, y = value, group = Period, color = Period)) +
# ylab("Abnormal Return, %") + xlab("Days relative to a volatility peak") +
# geom_line(size = 0.5) +
# geom_point(size = 2, shape = 21, fill = "white") +
# geom_hline(aes(yintercept = 0), linetype = "dashed") +
# scale_x_continuous(breaks = c(-10,-9,-8,-7,-6,-5,-4,-3,-2,-1,0,1,2,3,4,5,6,7,8,9,10)) +
# scale_y_continuous(breaks = c(-1,0,1)) +
# scale_color_discrete(labels = c("Total", "First", "Second")) +
# theme(legend.position="bottom")
#
# #Cumulative Abnormal Return over the event window. Total-, 1st- and 2nd period.

```

```

# Total.period = read.table("CAR.total.LT.txt",sep="\t", col.names=c("Total period"),
fill=FALSE, strip.white=TRUE)
# First.period = read.table("CAR.first.LT.txt",sep="\t", col.names=c("First period"),
fill=FALSE, strip.white=TRUE)
# Second.period = read.table("CAR.second.LT.txt",sep="\t", col.names=c("Second period"),
fill=FALSE, strip.white=TRUE)
#
# x <- -10:10 #days relative to a volatility peak
# y4 <- Total.period #total period
# y5 <- First.period #1st subperiod
# y6 <- Second.period #2nd subperiod
# df1 <- data.frame(x,y4,y5,y6)
#
# df3 <- melt(data = df1, variable.name = "Period", id.vars = "x")
#
# ggplot(data = df3, aes(x = x, y = value, group = Period, color = Period)) +
#   ylab("Cumulative Abnormal Return, %") + xlab("Days relative to a volatility peak") +
#   geom_line() +
#   geom_point(size = 2, shape = 21, fill = "white") +
#   geom_hline(aes(yintercept = 0), linetype = "dashed") +
#   scale_x_continuous(breaks = c(-10,-9,-8,-7,-6,-5,-4,-3,-2,-1,0,1,2,3,4,5,6,7,8,9,10)) +
#   scale_y_continuous(breaks = c(-4,-3,-2,-1,0)) +
#   scale_color_discrete(labels = c("Total", "First", "Second")) +
#   theme(legend.position="bottom")
#
# #AR - total period. Above and below median of amplitude
# x <- -10:10 #days relative to a volatility peak
# AR <- read.table("AR.total.LT.txt",sep="\t", col.names=c("AR"), fill=FALSE,
strip.white=TRUE)
# AR.above <- read.table("AR.above.LT.txt",sep="\t", col.names=c("Above"), fill=FALSE,
strip.white=TRUE)
# AR.below <- read.table("AR.below.LT.txt",sep="\t", col.names=c("Below"), fill=FALSE,
strip.white=TRUE)
#
# df4 <- data.frame(x,AR,AR.above,AR.below)
# df5 <- melt(data=df4, variable.name = "Amplitude", id.vars="x")
#
# ggplot(data = df5, aes(x = x, y = value, group = Amplitude, color = Amplitude)) +
#   ylab("Abnormal Return, %") + xlab("Days relative to a volatility peak") +
#   geom_line() +
#   geom_point(size = 2, shape = 21, fill = "white") +
#   geom_hline(aes(yintercept = 0), linetype = "dashed") +
#   scale_x_continuous(breaks = c(-10,-9,-8,-7,-6,-5,-4,-3,-2,-1,0,1,2,3,4,5,6,7,8,9,10)) +
#   scale_y_continuous(breaks = c(-6,-5,-4,-3,-2,-1,0,1,2)) +
#   scale_color_discrete(labels = c("All amplitudes", "Above median", "Below median")) +
#   theme(legend.position="bottom")

#Divide amplitude in above and below median of raising amplitude
ipeak2 <- vector()
for (i in 1:length(ipeak)){

```

```

  ipeak2[i] <- ipeak[i]
}
ipeak2 <- ipeak2[-1]
DFpeakamp <- data.frame(ipeak2,amplituder)
abind <- (rep(c(0,1),63))
abind <- c[1:126]
abind <- sort(abind) #order 0 and 1 as an indicator for below and above median amp.
new <- data.frame(DFpeakamp[ order(DFpeakamp$amplituder), ], abind)

ampbelow <- subset(new$Ipeak2,abind=="0")
ampabove <- subset(new$Ipeak2,abind=="1")

# FUGURES #
# ABOVE MEDIAN AND BELOW MEDIAN #
# #CAR - total period. Above and below median of amplitude
# x <- -10:10 #days relative to a volatility peak
# CAR <- read.table("CAR.total.LT.txt",sep="\t", col.names=c("CAR"), fill=FALSE,
strip.white=TRUE)
# CAR.above <- read.table("CAR.above.LT.txt",sep="\t", col.names=c("Above"),
fill=FALSE, strip.white=TRUE)
# CAR.below <- read.table("CAR.below.LT.txt",sep="\t", col.names=c("Below"),
fill=FALSE, strip.white=TRUE)
#
# df6 <- data.frame(x,CAR,CAR.above,CAR.below)
# df7 <- melt(data=df6, variable.name = "Amplitude", id.vars="x")
#
# ggplot(data = df7, aes(x = x, y = value, group = Amplitude, color = Amplitude)) +
# ylab("Cumulative Abnormal Return, %") + xlab("Days relative to a volatility peak") +
# geom_line() +
# geom_point(size = 2, shape = 21, fill = "white") +
# geom_hline(aes(yintercept = 0), linetype = "dashed") +
# scale_x_continuous(breaks = c(-10,-9,-8,-7,-6,-5,-4,-3,-2,-1,0,1,2,3,4,5,6,7,8,9,10)) +
# scale_y_continuous(breaks = c(-6,-5,-4,-3,-2,-1,0,1,2)) +
# scale_color_discrete(labels = c("All amplitudes", "Above median", "Below median")) +
# theme(legend.position="bottom")

```

11.1.7 Filename: Window.Table.Dummy.LT

```

#### THIRD STUDY ####
source("Algorithm-LT.R")
# Algorithm by Lunde and Timmermann (LT) #
#full,first and second period we use length=ipeak
#above and below median we use length=ampabove and length=ampbelow

window.tableR <- matrix(NA,nrow = length(ipeak),ncol = 11)
colnames(window.tableR) <- c("-10":"0")

window.tableR[,1] <- (ret-mean(ret))[ipeak-10]
window.tableR[,2] <- (ret-mean(ret))[ipeak-9]
window.tableR[,3] <- (ret-mean(ret))[ipeak-8]

```

```

window.tableR[,4] <- (ret-mean(ret))[ipeak-7]
window.tableR[,5] <- (ret-mean(ret))[ipeak-6]
window.tableR[,6] <- (ret-mean(ret))[ipeak-5]
window.tableR[,7] <- (ret-mean(ret))[ipeak-4]
window.tableR[,8] <- (ret-mean(ret))[ipeak-3]
window.tableR[,9] <- (ret-mean(ret))[ipeak-2]
window.tableR[,10] <- (ret-mean(ret))[ipeak-1]
window.tableR[,11] <- (ret-mean(ret))[ipeak]

window.tableF <- matrix(NA,nrow = length(ipeak),ncol = 10)
colnames(window.tableF) <- c("1":"10")

window.tableF[,1] <- (ret-mean(ret))[ipeak+1]
window.tableF[,2] <- (ret-mean(ret))[ipeak+2]
window.tableF[,3] <- (ret-mean(ret))[ipeak+3]
window.tableF[,4] <- (ret-mean(ret))[ipeak+4]
window.tableF[,5] <- (ret-mean(ret))[ipeak+5]
window.tableF[,6] <- (ret-mean(ret))[ipeak+6]
window.tableF[,7] <- (ret-mean(ret))[ipeak+7]
window.tableF[,8] <- (ret-mean(ret))[ipeak+8]
window.tableF[,9] <- (ret-mean(ret))[ipeak+9]
window.tableF[,10] <- (ret-mean(ret))[ipeak+10]

>window.window.bull <- c(bull[ipeak-10])

window.bull <- matrix(NA,nrow = length(ipeak),ncol = 11)
colnames(window.bull)<-c("-10":"0")
window.bull[,1] <- bull[ipeak-10]
window.bull[,2] <- bull[ipeak-9]
window.bull[,3] <- bull[ipeak-8]
window.bull[,4] <- bull[ipeak-7]
window.bull[,5] <- bull[ipeak-6]
window.bull[,6] <- bull[ipeak-5]
window.bull[,7] <- bull[ipeak-4]
window.bull[,8] <- bull[ipeak-3]
window.bull[,9] <- bull[ipeak-2]
window.bull[,10] <- bull[ipeak-1]
window.bull[,11] <- bull[ipeak]

window.bear <- matrix(NA,nrow = length(ipeak),ncol = 10)
colnames(window.bear)<-c("1":"10")

window.bear[,1] <- bull[ipeak+1]
window.bear[,2] <- bull[ipeak+2]
window.bear[,3] <- bull[ipeak+3]
window.bear[,4] <- bull[ipeak+4]
window.bear[,5] <- bull[ipeak+5]
window.bear[,6] <- bull[ipeak+6]
window.bear[,7] <- bull[ipeak+7]
window.bear[,8] <- bull[ipeak+8]

```

```
window.bear[,9] <- bull[ipeak+9]
window.bear[,10] <- bull[ipeak+10]
```

```
for (i in 1:length(ipeak)){
  for(j in 1:11){
    if(window.bull[i,j]==FALSE){
      window.tableR[i,j] <- NA
    }
  }
}
```

```
##LT full period
```

```
# window.tableR[56,1] <- NA #fix TURE values that are not close to the peak (FALSE
inbetween the TRUE and peak)
```

```
# window.tableR[56,2] <- NA
# window.tableR[71,1] <- NA
# window.tableR[71,2] <- NA
# window.tableR[71,3] <- NA
# window.tableR[71,4] <- NA
# window.tableR[71,5] <- NA
# window.tableR[71,6] <- NA
# window.tableR[71,7] <- NA
# window.tableR[72,1] <- NA
# window.tableR[91,1] <- NA
# window.tableR[91,2] <- NA
# window.tableR[97,1] <- NA
# window.tableR[97,2] <- NA
# window.tableR[100,4] <- NA
# window.tableR[100,5] <- NA
# window.tableR[121,1] <- NA
```

```
##LT first period
```

```
# window.tableR[56,1] <- NA
# window.tableR[56,2] <- NA
```

```
##LT second period
```

```
# window.tableR[13,1] <- NA
# window.tableR[13,2] <- NA
# window.tableR[13,3] <- NA
# window.tableR[13,4] <- NA
# window.tableR[13,5] <- NA
# window.tableR[13,6] <- NA
# window.tableR[13,7] <- NA
# window.tableR[14,1] <- NA
# window.tableR[33,1] <- NA
# window.tableR[33,2] <- NA
# window.tableR[39,1] <- NA
# window.tableR[39,2] <- NA
# window.tableR[42,4] <- NA
# window.tableR[42,5] <- NA
```

```

# window.tableR[63,1] <- NA

# #LT full period above median
# window.tableR[34,1] <- NA
# window.tableR[45,1] <- NA
# window.tableR[45,2] <- NA

# #LT full period below median
# window.tableR[11,4] <- NA
# window.tableR[11,5] <- NA
# window.tableR[29,1] <- NA
# window.tableR[30,1] <- NA
# window.tableR[30,2] <- NA
# window.tableR[59,1] <- NA
# window.tableR[59,2] <- NA

for (i in 1:length(ipeak)){
  for(j in 1:10){
    if(window.bear[i,j]==TRUE){
      window.tableF[i,j] <- NA
    }
  }
}

# #LT full period
# window.tableF[55,10] <- NA #Fix FALSE values that are not close to the (TRUE
inbetween the FALSE and peak)
# window.tableF[70,5] <- NA
# window.tableF[70,6] <- NA
# window.tableF[70,7] <- NA
# window.tableF[90,10] <- NA
# window.tableF[96,10] <- NA
# window.tableF[99,7] <- NA
# window.tableF[99,8] <- NA
# window.tableF[99,9] <- NA
# window.tableF[99,10] <- NA

# #LT first period
# window.tableF[55,10] <- NA

# #LT second period
# window.tableF[12,5] <- NA #Fix FALSE values that are not close to the (TRUE inbetween
the FALSE and peak)
# window.tableF[12,6] <- NA
# window.tableF[12,7] <- NA
# window.tableF[32,10] <- NA
# window.tableF[38,10] <- NA
# window.tableF[41,7] <- NA
# window.tableF[41,8] <- NA

```

```

# window.tableF[41,9] <- NA
# window.tableF[41,10] <- NA

# #LT full period above median
# window.tableF[39,5] <- NA
# window.tableF[39,6] <- NA
# window.tableF[39,7] <- NA
# window.tableF[58,10] <- NA
# window.tableF[59,10] <- NA
# window.tableF[61,10] <- NA

# #LT full period below median
# window.tableF[48,7] <- NA
# window.tableF[48,8] <- NA
# window.tableF[48,9] <- NA
# window.tableF[48,10] <- NA

window.table <- matrix(NA,nrow = length(ipeak),ncol = 21)
colnames(window.table) <- c("-10":"10")

window.table[,1:11] <- window.tableR
window.table[,12:21] <- window.tableF

AR.total.LT <- apply(window.table,2,mean,na.rm=TRUE)*100 #abnormal return of all peaks
for n-days before and after
AR.total.LT
t.test(AR.total.LT)
CAR.total.LT <- cumsum(AR.total.F) #cumulative abnormal return
CAR.total.LT
t.test(CAR.total.LT)

#save to a file
# write.zoo(AR.total.F, file="AR.total.LT.txt") #total = full period
# write.zoo(CAR.total.F, file="CAR.total.LT.txt")
# write.zoo(AR.first.F, file="AR.first.LT.txt") #first = 1st half
# write.zoo(CAR.first.F, file="CAR.first.LT.txt")
# write.zoo(AR.second.F, file="AR.second.LT.txt") #second = 2nd half
# write.zoo(CAR.second.F, file="CAR.second.LT.txt")
# write.zoo(AR.above.F, file="AR.above.LT.txt") #above median
# write.zoo(CAR.above.F, file="CAR.above.LT.txt")
# write.zoo(AR.below.F, file="AR.below.LT.txt") #below median
# write.zoo(CAR.below.F, file="CAR.below.LT.txt")

```

11.2 Reflection Note

In the thesis we look at stock return in relation to measured volatility. The measured indices are S&P 500 and VIX respectively. Both indices are benchmark for international finance influencing stock exchanges all over the world. Moreover, the thesis conducts three studies, one of the specific relationship between S&P 500 and VIX. One studying the VIX index in terms of rising and falling volatility-phases. And at least, one study of abnormal return around the spike peaks in the VIX index. In the literature investors sentiment are described as an amplifying force of determine volatility. The findings clearly state a negative relationship between VIX and S&P 500. When studying the phases in volatility, the findings show a longer duration for falling phases when rising phases increase more than falling phases decreases. Which can be interpret as the stock return increases for longer periods and drop faster, due to the negative relationship between S&P 500 and VIX. The abnormal return around the peaks are significantly negative before and including peak day, and significantly positive after peak day. Which mean that the peak of the VIX index can be an indicator to buy. To outline the studies in the thesis, knowledge from courses in the master program has been necessary. Especially the financial and econometrical courses as Finance Theory, Econometric for Finance and Computational Finance.

The school of Business and Law at the University of Agder consider internationalization, innovation and responsibility as core areas within the the field of business. Throughout our master program the internationalization has been an embedded part of our learning. The fact that every course is taught in English make us ready for an international working life, where English is more and more acceptable as the day to day language, even in companies dealing with business in Norway. As of our thesis the view is highly international with an international theme including the biggest index in financial markets in the world. The findings are adaptable to all other benchmark indices around the world. Even a singular company listed on a stock exchange are in the relevance of the findings, though the risk is not diversified. New ways of buying stocks or investing in foreign markets are in constant development. This make the results from the US market more relevant to Norwegian investors than a few years ago. It seems like there is no limit to the globalization we see today, and the international relevance of the thesis is higher since we studied the US market instead of the Oslo Stock Exchange.

As of the innovation related to the theme we conduct two new studies. The studies relate to algorithms of finding bull and bear market, in the thesis the algorithm are used in a different context, determine phases of the volatility index, VIX. Further the study of abnormal return is not find in the literature when using phases of rising and falling volatility, still we do not consider the study to be innovative since the relationship is already observed. For innovation to be more relevant for the theme, future studies have to be in place. Where it is possible to write new algorithms for finding a turning point on the date of the turning point, but more research and new studies have to be outlined. Algorithms for investment fund exist today where robots buy and sell stock by themselves. To include our findings in an algorithm in a simulating market would be interesting, when take the results a step further the innovation is more relevant. All though, international financial markets are a driving force for internationalization through its innovating financial instruments and systems for buying stock in foreign countries online. From our data set we see at large increase in VIX and decrease in S&P 500 as a part of the financial crash during fall 2008. The crisis included a large amount of innovative and creative financial instrument at the time, derivatives such as CDO's and CDS's.

As mention above the theme of the thesis uses two indices in the analysis. In example the S&P 500 index are a combination of the 500 largest stocks in US. The market place responsible for the indices have a responsibility to be fair and transparent for all investors. In financial markets there are many advanced investment opportunities, such as derivatives, options, futures, forwards or a combination of these. When dealing with advanced financial products the advisors have a responsibility of act in the best interest for the client. Advisors can have other incentives to do a transaction on behalf of the client, for example own gain as commission. If we are in a position as advisors or employees of a market place in the future, the need for high integrity and moral have to be a part of our work. Throughout our master program we have learned about theory in financial markets and different strategies in business. The theory assumes perfect markets which are not the fact in reality. Responsibility and integrity in own work especially in finance is important, the financial markets can provide great return on investment, as well as significant losses. Responsibility is not to be too greedy and not take any shortcuts that may be illegal.

As the analysis of the thesis look at the relation of two market indices, one for stocks and one for volatility, the potential ethical challenges that may arise are ethical challenges related to financial markets. Some of the challenges are mention as the reflection of responsibility, as well

as dealing with sensitive information about the stock market (as sensitive information about a stock listed at a stock exchange) or be in a position able to act on inside trading.