Norges Handelshøyskole

FIN11 Trading and Market Microstructure

Term paper

Predefined Information

Start date:	2020-11-03 09:00 AM	**Term:**	2020 12
End date:	2020-12-14 02:15 PM	**Assessment type:**	Norwegian 6-point scale (A-F)
Examination type:	Term paper (70%)		
SIS Code:	191 FIN11 1 O3 2020 12		
Internal assessor:	(Anonymised)		

Participant

Candidate No.:	110039

Group

Group name:	(Anonymised)
Group number:	7
Other members:	110053, 110015, 110010

DECEMBER 14, 2020

ALGORITHMIC TRADING

DEVELOPING, ANALYZING AND OPTIMIZING A TRADING ALGORITHM



Figure 1 (Smartereum, 2019)

CANDIDATES: 110010, 110015, 110039, 110053

TERM PAPER
TRADING AND MARKET MICROSTRUCTURE

WORD COUNT: 6342

Abstract

Over the past 20 years, financial markets in most regions of the world have experienced growth in the use of algorithmic trading across different asset classes (Goldman Sachs, 2018). This new way of trading has caught a lot of attention, including ours. Therefore, this paper aims to shed light on the subject of algorithmic trading. First, this paper introduces the concept of algorithmic trading as well as addressing both its advantages and disadvantages, mostly related to the reduction of human intervention. Thereafter, some important key concepts and theories are presented, including asset types, volatility, transaction costs, as well as performance measures.

In the next part of this paper, we explore the process of designing a trading algorithm with a practical example. We begin with presenting the strategy and translating it into code. The idea is that one can profit from the difference between realized and implied volatility. The first strategy uses the GARCH model – a statistical model often used to measure realized volatility in financial markets. Meanwhile the second strategy uses the standard deviation of an S&P500 ETF, the SPY, instead of the GARCH model. Further, the volatility index for the S&P500 index – the VIX – is used as implied volatility. The strategy is simple: to short volatility when realized volatility is lower than implied, and to go long when realized volatility is higher than implied, which means that we will be invested at all times. The strategy using the GARCH model, the strategy using the standard deviation of SPY and the average of these strategies is then backtested with SPY as the benchmark, in order to analyse their performance. Further, the algorithms are optimized to reduce the implied risk.

What we have found is that the average of the two strategies clearly outperforms the buy and hold strategy with SPY. The strategies outperform not only in terms of higher returns, but also in all of the three performance measures: Sharp-Ratio, Calmar-Ratio and Ulcer Performance-Index. However, these benefits come at a cost. The strategies increase the transaction costs due to higher frequency of trading. Further, the strategies also result in a higher standard deviation, and higher risk as the algorithm walks right into the volatility crash of 2018. In order to prevent this, we optimized the algorithm, by incorporating the concept of decreasing the invested position when profits increased. Following the optimization, the algorithm no longer walks into the volatility crash of 2018, giving reduced risk and reduced standard deviation. Although this also means a significant reduction in returns, the algorithm still greatly outperforms SPY.

Contents

Abstract 1
1 Introduction – A New Age of Quantitative Trading 3
2 Theory 5
2.1 Algorithmic Trading – Advantages and Disadvantages 5
2.2 ETF – Exchange Traded Fund 6
2.3 ETN – Exchange Traded Note 6
2.4 Volatility 6
2.5 Transaction- and Management Costs 7
2.6 Sharpe Ratio 7
2.7 The Ulcer Performance Index 8
2.8 The Calmar Ratio 9
3 Data and methodology 10
3.1 Strategy and Model Building 10
3.1.1 Strategy 1: GARCH Predictions as Realized Volatility 10
3.1.2 Strategy 2: Standard Deviation as Realized Volatility 12
3.2 Backtesting 13
4 Results 16
4.1 Performance Analysis 16
4.1.1 Returns and Drawdowns 16
4.1.2 Sharpe, Calmar, and Ulcer 18
4.1.3 Transaction- and Management Costs 20
4.2 Optimization of the Algorithm 21
4.2.1 Modifying the Algorithm 21
4.2.2 Returns and Drawdowns 24
4.2.3 Sharpe, Calmar, and Ulcer 25
4.2.4 Transaction- and Management Costs 27
5 Conclusion 28
References 29

1 Introduction – A New Age of Quantitative Trading

As the computer has become smarter and more widespread than ever, they are being used to support humans in all imaginable areas of life and business. In finance, algorithmic trading has become an integral and permanent part of our capital markets (SEC, 2020). That has led to a great increase in market liquidity, decrease in volatility, reduced trading costs, and made the markets more accessible than ever (Harris, 2003). However, the computers have also been blamed for causing several of the most extreme market crashes in the history of the stock market (SEC & CFTC, 2010).

Since the turn of the century, algorithmic trading has "disrupted the structure of the market and now makes up a significant amount of the daily trading volume on US and Global Exchanges" (Dehod, 2019). Within the money management industry, we observe an ever-increasing use of quantitative method and computer programs to automate their highly sophisticated strategies. At least 5 of the top 10 largest hedge funds worldwide (based on AUM) can be regarded as being quantitative, in that they approach the market with mathematical, computational, and statistical methods (Bloomenthal, 2020). The question that arises is: "are the quant-strategies really that good?". The performance of Renaissance Technologies' flagship fund – Medallion, speaks for itself:

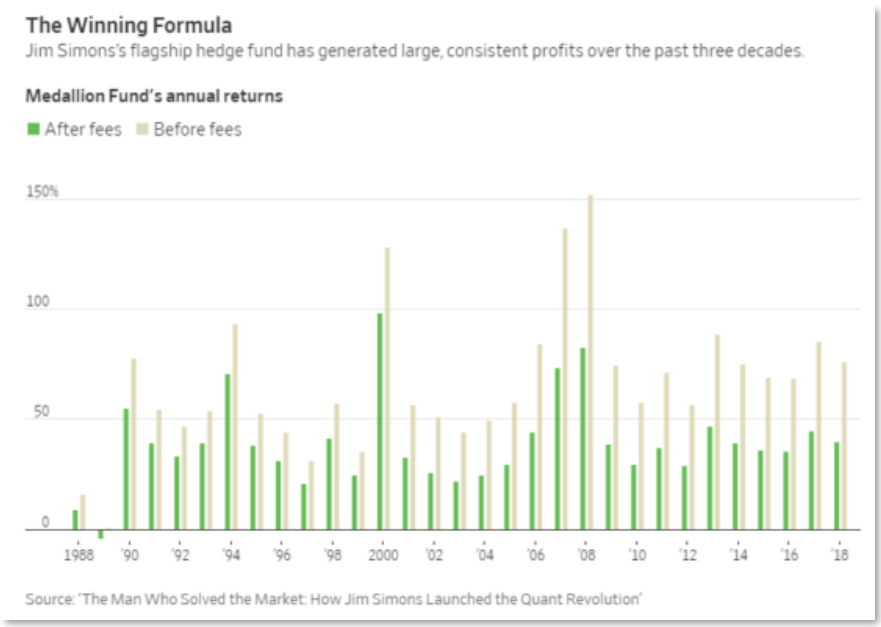


Figure 2: Medallion Fund Returns

The fund has delivered a staggering 39 percent annual return after one of the industry's highest fees, over 30 years (Remsburg, 2020)! The founder, a mathematician and former cold war code breaker named Jim Simons, has recently been regarded as the world's greatest investor (Zuckerman, 2019). Just 20-30 years ago, most people were sceptical about this new breed of math whizzes starting to apply their technical skills to "solve the markets". However, now they seem to have conquered Wall Street (Patterson, 2011). No wonder why more and more people, both institutional- and retail investors, are starting to explore the world of algorithmic and quantitative trading (SEC, 2020).

We have found algorithmic trading to be one of the most fascinating areas in finance and have therefore decided to concentrate our paper around the process of developing a trading algo. More specifically, we will explore how algorithmic trading affects returns and risk, and how we through optimization can adjust the strategy to reduce risk. The results will be discussed and compared with a passive buy and hold strategy.

Firstly, we will present a few key concepts and theories that will be of use throughout our paper. Secondly, the process of building and backtesting a trading algo will be elaborated upon. Here it is important to emphasise the word "process". We have no intention of building the next Renaissance Technologies, but rather analyse and optimize existing strategies using multiple performance measures. Finally, we will provide a conclusion based on our findings and address potential weaknesses.

2 Theory

We will begin this section by discussing the concept of algorithmic trading. What is it and what advantages and disadvantages does it represent? Next, the instruments used in the strategies will be explained. Thereafter, volatility and transaction costs will be presented, both of which may affect the performance of the strategy. Finally, three performance measures will be introduced.

2.1 Algorithmic Trading – Advantages and Disadvantages

Algorithmic trading is defined by Investopedia as:

"*[…] a computer program that follows a defined set of instructions (an algorithm) to place a trade. The trade, in theory, can generate profits at a speed and frequency that is impossible for a human trader"* (Seth, 2020).

In comparison to more traditional ways of trading, algorithmic trading has both advantages and disadvantages (Samuelsson, 2020). Starting off with its advantages, algorithmic trading offers both increased speed as well as more accuracy. These advantages come as a result of computers replacing human intervention. Further, algorithmic trading takes away the emotional aspect of trading. For instance, an algorithm will not sell in panic if a stock drops a bit more than anticipated. It will not keep a stock just because it likes the company either. The algorithm does exactly what it is programmed to do, which forces the trader to be 100 percent specific in his goals and ideas. At last, algorithmic trading offers the ability to backtest your strategy using historical data, which might help improve and optimize your strategy (Samuelsson, 2020).

Next, moving on to disadvantages of algorithmic trading. First of all, the loss of human control also means lack of human sense. Algorithms cannot completely understand the market situation and all circumstances in the same way as humans. Further, algorithmic trading also means dependency of technology. The technology must be monitored to make sure it works as intended, if not there is a risk that you end up losing money (Samuelsson, 2020). In addition, a strategy that works today, might not work tomorrow.

2.2 ETF – Exchange Traded Fund

According to ETFGI, a research and consulting company within the ETF-industry, the amount of cash invested in ETF´s during the last ten years is more (ETFGI, 2019) than fivefold. But what exactly is an ETF? An Exchange Traded Fund is a fund that can consist of multiple securities which at all time can be traded at an exchange. For instance, it can consist of stocks, indexes, commodities, and bonds. Most other funds are traded once during a day, while ETF´s are traded just like stocks throughout the day (Chen, Exchange Traded Funds - ETFs, 2020). What this means is that ETF´s offers both the flexibility of a stock and the diversification from funds, in a single financial instrument. Further, it is also less costly to buy an ETF than to buy all of the underlying's.

2.3 ETN – Exchange Traded Note

ETN, or Exchange Traded Note, is an unsecured debt security issued by financial institutions. Just like ETF´s, ETN´s are traded at an exchange like stocks throughout the day, which means that it offers flexibility. ETN´s can most easily be understood like a bond. It does not pay interest payments, but at maturity it aims to pay returns according to its underlying index. Contrasting to ETF´s, ETN´s does not provide ownership in the underlying but is backed by the credit of the issuer. This means that the risk in ETN´s is not only the performance of the underlying, but also if the issuer is able to repay at maturity (Chen, Exchange Traded Notes - ETN, 2020).

2.4 Volatility

Volatility is the unexpected tendency for changes in prices. The change can come as a response to new information about value, or as a response to a demand from traders asking liquidity, i.e., the ability to trade large volume without it being too costly (Harris, 2003). Further, volatility is closely related to both risk and profit. If the price of the underlying drops, it means that you on paper lose money if you are long, or earn money if you are short, and vice versa. Because of changes in price, volatility influence risk. An underlying that fluctuates a lot is consequently riskier than a steadier underlying (Harris, 2003).

Furthermore, volatility can be divided into two types: Fundamental volatility and Transitory volatility. Fundamental volatility is a result of unanticipated changes in instrument values.

Factors that might impact the instrument value, are for instance the quality of the management, supply and demand for the product offered, the company´s resources and technology, interest- and inflation rates as well as macroeconomic policies (Harris, 2003). Transitory volatility is a result of trading activity by impatient uninformed traders. This type of volatility causes the price to diverge from fundamental value but is named transitory because the prices eventually will revert to fundamental value. Most of the time this type of volatility is due to what is called "*Bid/Ask Bounce*", which happens when market orders are used to buy at ask or sell at bid (Harris, 2003). Adding both the fundamental- and transitory-volatility gives the Total volatility, which is used as a measure of volatility by its variance, standard deviation or mean absolute deviation of the change in price (Harris, 2003).

2.5 Transaction- and Management Costs

Transaction costs represents costs associated with trading and are divided into three types (Harris, 2003). The first type is "*Explicit Transaction Cost*", which are the costs that are most easy to identify. Taxes to the government and management costs like commissions are good examples. The second type is "*Implicit Transaction Cost*", which are the costs that follows from trader's impact on prices. If you use a market order when there is a large difference between the bid and the ask, you pay what is called the spread. Further, when execution larger orders, one might run the risk om walking up the book. The bigger the order, the larger the impact. These are typically implicit costs. The last type is the "*Missed Trade Opportunity Cost*" or time cost. A good example is if you want to buy a stock and place a limit-order around the lowest ask. If the price then rises, and the order fails to execute, you have lost the price-movement (Harris, 2003).

2.6 Sharpe Ratio

Numerous performance measures can be used to evaluate the performance of any trading strategy or financial security. One of the most known performance measures is the Sharpe ratio. The measure was developed by William F. Sharpe in 1966 and adjusts the return of an investment with the risk taken (Hargrave, 2020). The measure is calculated by taking the difference between the returns of the investment and the risk-free rate and dividing it by the standard deviation of the investment. Because an investor would ideally want high return with close to no risk, higher is better for the Sharpe.

$$Sharpe\ Ratio = \frac{R_p - R_f}{\sigma_p}$$

$R_p = return\ on\ investment$

$R_f = risk - free\ rate$

$\sigma_p = standard\ deviation\ of\ the\ investment$

U.S. Treasury rate is commonly used as the risk-free rate, although it theoretically depends on the U.S. government's ability to tax its citizens (Chen, Treasury Bond (T-Bond), 2020).

2.7 The Ulcer Performance Index

When using the Sharpe Ratio, we are threating the volatility of the market as being normally distributed. However, history has shown that extreme down movements are usually more extreme than extreme up movements, meaning the return distribution has a fat left tail. Instead of using the standard deviation, one can use other volatility measures that do not assume normality. One such measure was created by Peter Martin in 1987. The measure is called the Ulcer index and tells us something about the depth and length of drawdowns of the performance. For every new day in a specified period, the index calculates an R-score. The R-score represent the percentage drawdown from the previous highest price. Thus, the "maxprice" is updated for every new high.

$$R_i = 100 * \frac{price_i - maxprice}{maxprice}$$

Further, these R-scores are squared, divided by the number of days studied, and then taken the square root of to produce the Ulcer index.

$$Ulcer\ index = \sqrt{\frac{R_1^2 + R_2^2 + \cdots + R_N^2}{N}}$$

The Ulcer index is often used as a technical indicator by technical traders. It will tell the trader something about the current volatility in the market. However, it can also be used as a

replacement for the standard deviation in the Sharpe ratio formula. The ratio has been given its own name: The Ulcer Performance Index or UPI. This is the performance measure that we will be applying in our analysis. In contrast with the standard deviation, the Ulcer index will exaggerate lasting drawdowns because for every new day, a new R squared will be added to the index and thus blowing up the number. The UPI should be interpreted like the Sharpe. Hence, the higher the better.

$$UPI = \frac{R_p - R_f}{Ulcer\ index}$$

2.8 The Calmar Ratio

The Calmar ratio, or sometimes called the drawdown ratio, was created by Terry W. Young and published in 1991. Like the Ulcer performance index, the Calmar ratio focuses on the drawdowns. Young argued that no investor worry about the upside risk, they all worry about the downside risk. The ratio measures the average annual rate of return of an investment compared with its maximum drawdown (Kenton, 2020). It is calculated by dividing the average annual rate of return by the maximum drawdown (in percent) in the period. Higher is better.

$$Calmar\ Ratio = \frac{Average\ annual\ rate\ of\ return}{Maximum\ drawdown}$$

3 Data and methodology

In this section, we will introduce two trading strategies and convert them into code. The strategies will serve as the foundation for the rest of our paper. Translating the strategies into computer programs allows us to backtest the strategies, obtain performance metrics, and measure the performance of the algorithms.

3.1 Strategy and Model Building

Both strategies will be based on trading VIX – a volatility index which represents the market's expectations of volatility over the next 30 days, based on the implied volatility of options for the S&P500 index. These strategies use what in described as Volatility Risk Premium in Tony Cooper's seminal paper "Easy Volatility Investing" (Cooper, 2013). The idea is that you compare a measure of realized volatility with implied volatility, i.e., with VIX. When realized volatility is higher than implied, it means the market is in a panic and we want to be long in VIX as volatility will rise. Contrarily, when realized volatility is lower than implied, things are as they should be, and we can then collect volatility premium by shorting the VIX.

3.1.1 Strategy 1: GARCH Predictions as Realized Volatility

For the first strategy we will use a GARCH model to calculate realized volatility. GARCH is short for Generalized Auto-Regressive Conditional Heteroskedasticity. I.e., a generalized form of using past values to predict future values, that the present value differs given a past value, and lastly that volatility varies over time (Kipnis, 2018). Summarized, this amounts to using past volatility to predict future volatility given that it changes over time.

The use of GARCH in analyses of time-series data assumes two things. Firstly, the variance error is believed to be serially autocorrelated, meaning that we assume history repeats itself or the correlation of a signal with a delayed copy of itself, i.e., a similarity between a time series and a lagged version of itself over successive time intervals (Smith, 2020). In practice this could mean that a security's value rises and fall in a repetitive pattern, as seen in the value of VIX (Figure 3).

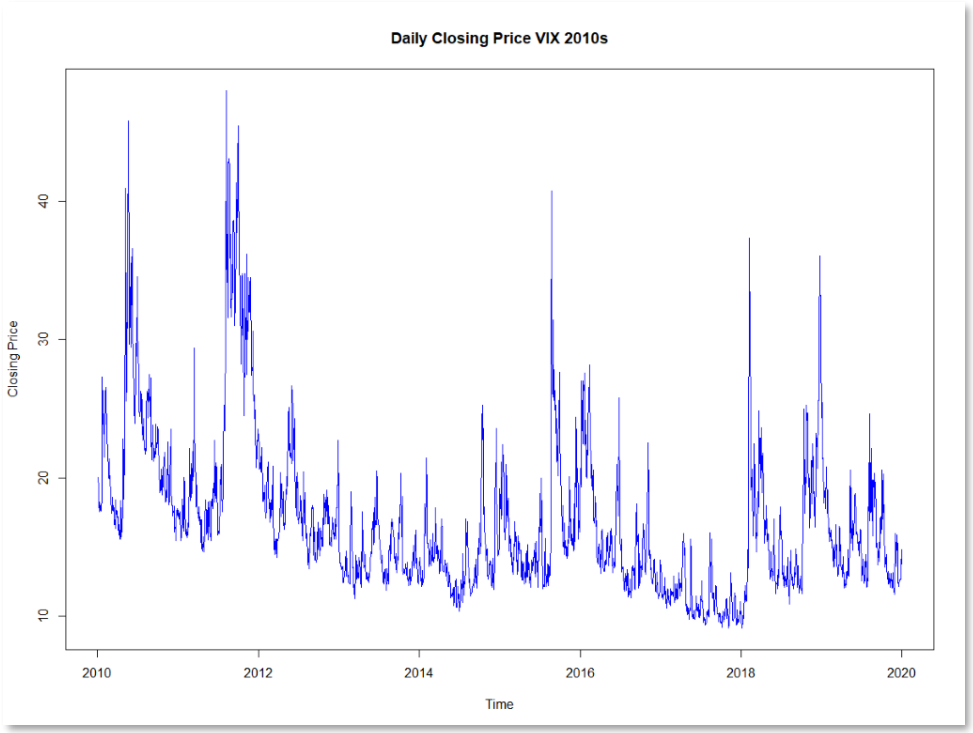


Figure 3: VIX (2010-2020)

Secondly, it assumes that the variance of the error term follows an autoregressive moving average (ARMA) process, where autoregressive denotes regressing a variable on its own past data and the moving average means that there is a linear combination of error terms occurring both contemporaneously as well as at various times in the past (Smith, 2020). An AR(1) model is used in the strategy, which specifies to the program that for the GARCH predictions, the VIX follows a random walk (Quantstart, 2020).

Using GARCH is most appropriate in time series where the variance of the error term is serially autocorrelated following an autoregressive moving average process. It is, therefore, most used for assessing risk and expected returns for securities that experience clustered periods of volatility in return (Kipnis, 2018). From Figure 3 and Figure 4 we can see that this is the case with VIX –volatility clusters in tall spikes for extended periods, contrary to single-day spikes and then calm days following directly. From these figures we can see that VIX is reverting, in that when it reaches a top it does not stay elevated but rather declines after some time to a previous value. This supports our hypothesis that using GARCH as a measure of realized volatility and VIX as a measure of implied volatility could have a potential for being a profitable trading strategy.

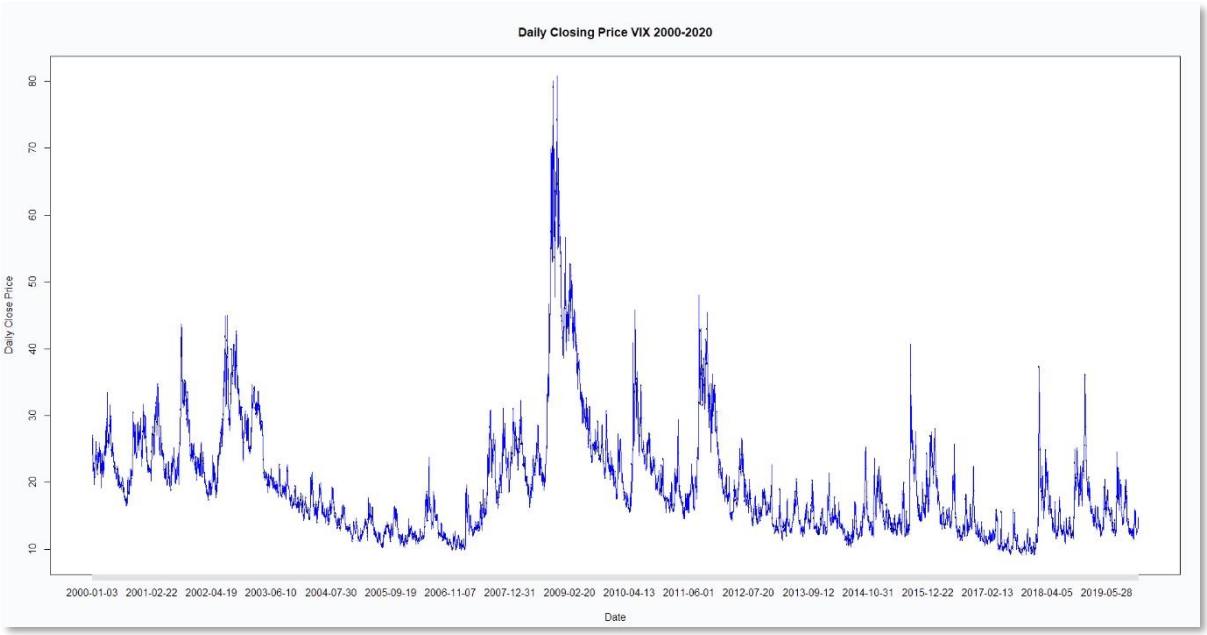


Figure 4: VIX (2000-2020)

3.1.2 Strategy 2: Standard Deviation as Realized Volatility

For the second strategy VIX will be kept as a measure of implied volatility while the measure of realized volatility will be changed from GARCH to the standard deviation of an S&P500 ETF – the SPY. The reasoning for why the standard deviation of SPY is a good alternative measure of realized volatility is as follows: the previously used GARCH model has three assumptions (Kipnis, 2018). Firstly, the mean model, i.e., an assumption about the ARMA in our strategy is set equal to AR(1). This means that we assume that the VIX follows a random walk, which means that its value cannot be predicted. Secondly, the variance model, the type of GARCH model and variance targeting forces an assumption of mean reversion. Finally, an assumption about the distribution model of the returns, and that these returns can be approximated to a Skewed Student T Distribution. When looking at Figure 5, which shows a chart for the S&P500 which SPY aims to replicate, we can see that positive returns usually come as the sum of small gains, whereas negative returns are large chunks, as we have seen in 2020. This fits for the profile of a Skewed Student T Distribution, which means using the standard deviation of SPY as a measure of realized volatility is a reasonable alternative.

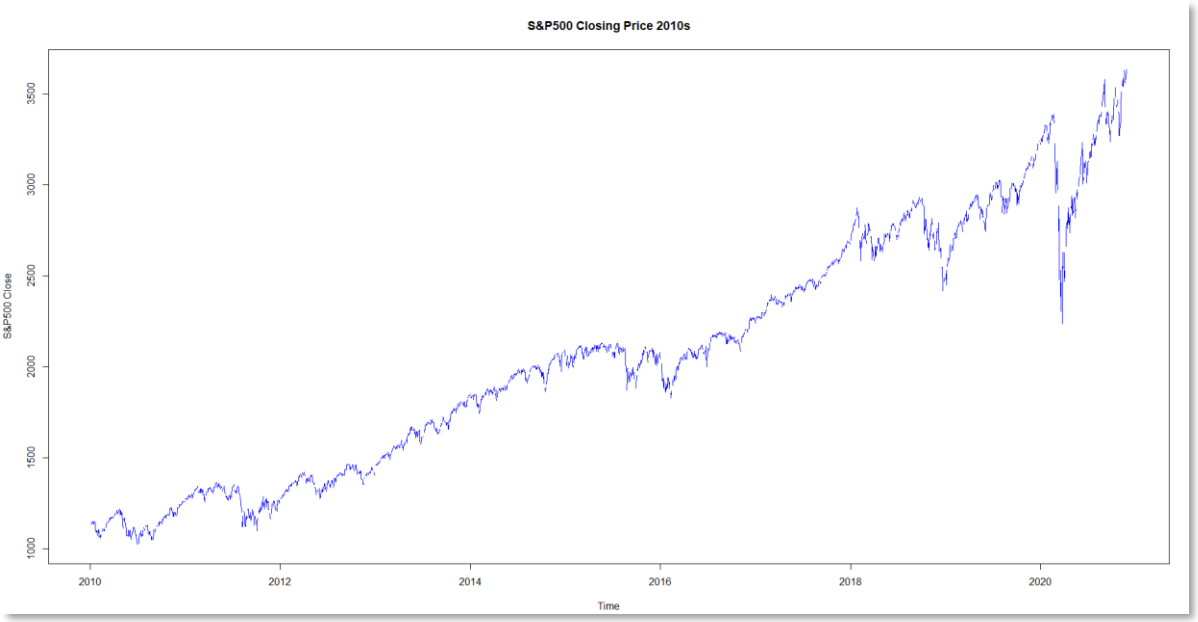


Figure 5: S&P500 (2010-2020)

3.2 Backtesting

Backtesting is an important step in the process of developing a trading algorithm, as it reveals how the strategy would have performed historically. To backtest the two strategies, we need to translate them into code. The implementation of these strategies into code was developed by Ilya Kipnis, one of the highest-ranking individuals in the open-source R/Finance community and published on the web page "QuantStrat TradeR" (Kipnis, 2018). The implementation begins by specifying the three components of algorithm that calculates the GARCH predictions for the first strategy.

```
gjrSpec = ugarchspec(mean.model = list(armaOrder = c(1,0)),
                     variance.model = list(model = "gjrGARCH",
                                           variance.targeting = TRUE),
                     distribution.model = "sstd")
```

Figure 6: GARCH model

The mean model, variance model and distribution model are set as specified in the strategy with AR(1), GARCH model with variance targeting that assumes a degree of mean reversion and a skewed student T distribution. After creating the model for predicting volatility, we need to create a rolling out-of-sample prediction. Important parameters to pass to the computer are the

number of days in the rolling window and refitting of the model. The strategy uses 504 days and 22 days, or approximately a trading month.

```
garchroll = ugarchroll(gjrSpec, data = spyRets,
                       n.start = 504, refit.window = "moving", refit.every = 22)
```

Figure 7: Model specifications

We could have chosen to refit the model more often, though this would put extra strain on the computer; it took three minutes to compute with refitting every 22 days. Any interval shorter than that would not be beneficial due to time strain. In addition, one could also run into fail-to-converge instances, that would disallow converting predictions to a data frame. The result is shown in Figure 8.

```
> garchroll
                   Mu       Sigma      Skew    Shape Shape(GIG)      Realized
1995-01-30  6.649495e-06 0.005553869 0.9456196 4.116597          0 -0.0043116581
1995-01-31  4.946583e-04 0.005635317 0.9456196 4.116597          0  0.0039968147
1995-02-01  6.653709e-06 0.005592602 0.9456196 4.116597          0 -0.0003314729
1995-02-02  2.608790e-04 0.005555797 0.9456196 4.116597          0  0.0059735226
1995-02-03 -1.094498e-04 0.005522803 0.9456196 4.116597          0  0.0141863067
1995-02-06 -5.918340e-04 0.005493878 0.9456196 4.116597          0  0.0042297746
```

Figure 8: Volatility predictions

The sigma quantity, as it is the volatility predictor, is the most important and is the quantity the program will compare to the VIX. Since it is not possible to short the VIX, we will use two instruments to trade this index. When we want to short VIX, we will open a long position in ZIV – an ETN which has the aim of producing inverse results of VIX. On the other hand, when we want a long position in VIX, we will open a long position in VXZ – an ETN which seeks to directly reflect the performance of VIX. It must be noted that ZIV was delisted on July 12th, 2020 (Credit Suisse, 2020). Therefore, if either of the two strategies was to be used for trading currently, then one could either directly short VXZ, or wait for a proposed new ETF – SVIX, which would have the aim of being inverse of VIX, similarly to ZIV (Kawa, 2019).

To make the program trade when the strategy calls for it, we define a variable “diff” that is the difference between GARCH sigma predictions annualized and multiplied by 100 and VIX

predictions for the first strategy. For the second strategy, the "diff" variable is calculated as the difference between the standard deviation of SPY and VIX.

When the variable diff has a positive value, i.e., realized volatility is greater than implied volatility, the algorithm creates a signal to go long volatility, which is vxzSig in the code. This means the algorithm will be buying shares of the VXZ ETN. In other words, we are betting on volatility to increase in the future, based on the fact that traders are more relaxed.

Meanwhile, if the variable diff has a negative value, that is, when implied volatility is greater than realized volatility, the algorithm creates a signal to go short volatility, which is zivSig in the code. This means the algorithm will be buying shares of the ZIV ETN.

```
zivSig = diff < 0
vxzSig = diff > 0

garchOut = lag(zivSig, 2) * zivRets + lag(vxzSig, 2) * vxzRets

histSpy = runSD(spyRets, n = 21, sample = FALSE) * sqrt(252) * 100
spyDiff = histSpy - Ad(VIX)

zivSig = spyDiff < 0
vxzSig = spyDiff > 0

spyOut = lag(zivSig, 2) * zivRets + lag(vxzSig, 2) * vxzRets

avg = (garchOut + spyOut)/2
```

Figure 9: Signals

4 Results

In the following section, the performance of the strategies will be analysed using the performance measures introduced in the theory section. Further, we will seek to optimize the strategies to reduce risk.

4.1 Performance Analysis

To get a better measure of the quality of the trading algorithm, one should compare its performance to a benchmark. In this case, we will use SPY as the benchmark, as it represents the best alternative to our strategy, i.e., passively holding a fund which tracks the performance of the S&P500. Another reason for choosing SPY is us having chosen it earlier to calculate historic volatility for our second algorithm. Furthermore, the performance metrics that we will use for the performance analysis are cumulative return, worst drawdown, annualized return, standard deviation, and Sharpe Ratio. In addition, the Calmar ratio and Ulcer Performance index are used.

4.1.1 Returns and Drawdowns

Figure 10 below shows the cumulative return of the two algorithm strategies as presented in the hypothesis section, the average of those two strategies, and the returns of SPY between 2004 and 2019. The returns of both algorithms are close, however, during some periods, notably in 2009-2010 and 2015, they varied substantially. We can also see that both strategies and the average of the two perform substantially better than passively holding SPY.

The drawdown of the algorithms and SPY are also plotted. These show periods of substantial drawdown of over -50%, meaning the algorithms lost over half of their value during certain periods of time. This clearly visualizes the amount of downside risk the algorithms are exposed to during periods of high volatility. Apart from this, drawdown rarely moves below or near -40%, except for in 2016 where drawdown dropped for the algorithm which uses GARCH.

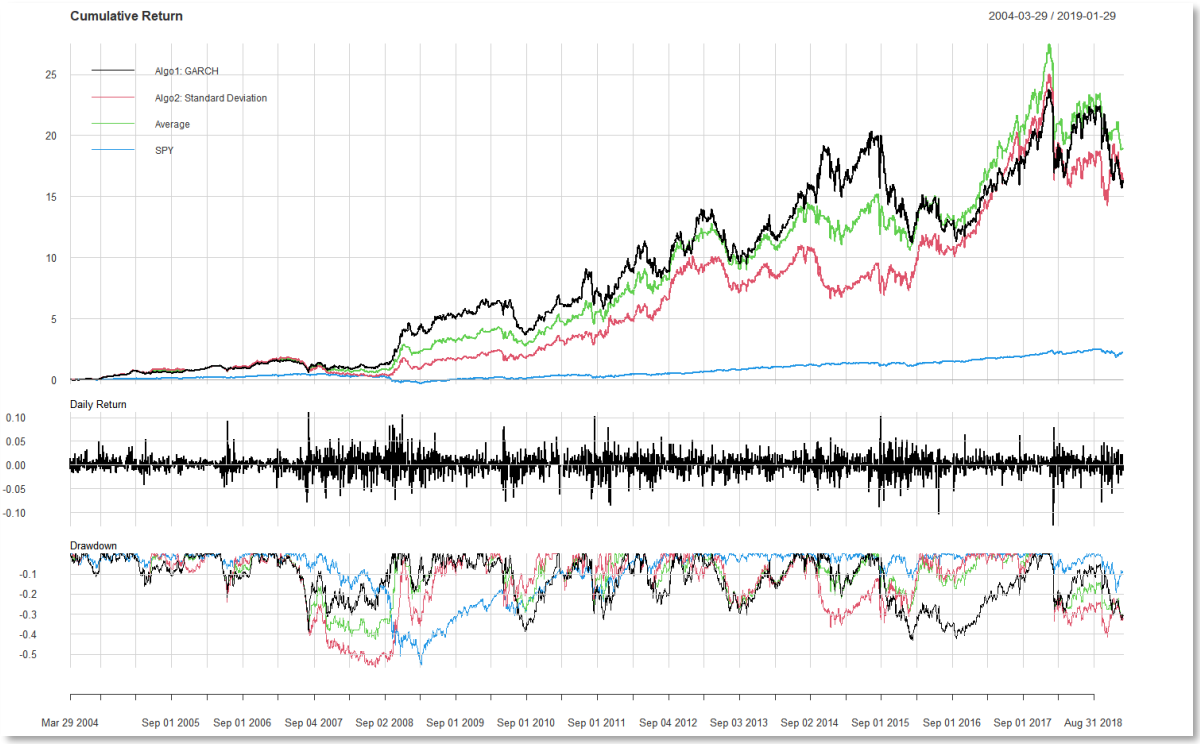
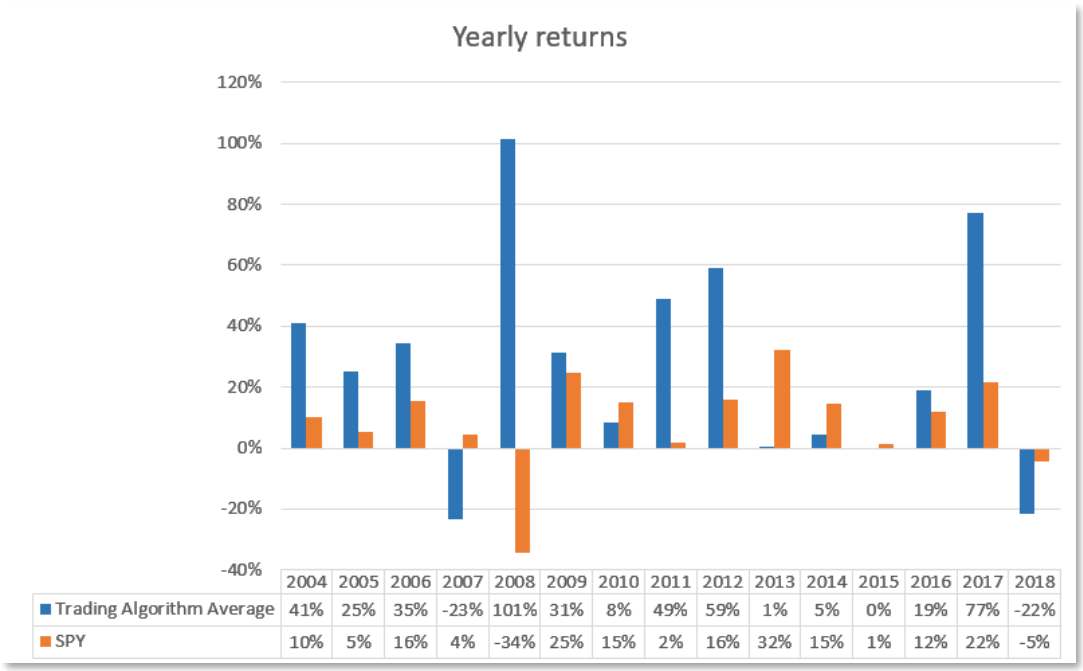


Figure 10: Cumulative Return - Initial model

To further explore the behaviour of the algorithms, we can look at Figure 11 which illustrates the year-by-year returns of the average of the two strategies plotted against the returns of SPY. The figure once again confirms that the average of the two strategies generally gives a higher yearly return than passively holding SPY. We can also observe that the average and SPY are most different during years with rapid change of volatility: the build-up to the Great Recession in 2007 and the Great Recession itself in 2008, as well as the build-up to the Volatility Crash in 2017 and the Volatility Crash itself in 2018. While 2008 and 2017 give favourable results with very significant returns when using the algorithm, 2007 and 2018 present notable drawdowns of over 20%. This underlines that there is potential for improving the algorithm.



	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Trading Algorithm Average	41%	25%	35%	-23%	101%	31%	8%	49%	59%	1%	5%	0%	19%	77%	-22%
SPY	10%	5%	16%	4%	-34%	25%	15%	2%	16%	32%	15%	1%	12%	22%	-5%

Figure 11: Yearly Returns - Average of the initial models

Although these figures give a good indication of the algorithms' performance, it is necessary to obtain more precise measures on the latter. This can be done by looking at the table in Figure 11 which summarises the results from the backtest. We can see that both algorithms give an approximately equal annualized return at 21%, beating the annualized return from holding SPY, which lies at 8%. We should also notice that the average of the two algorithms yields a higher annualized return at 22%, which highlights the importance of diversification for algorithmic trading.

4.1.2 Sharpe, Calmar, and Ulcer

While the two algorithms give higher returns than SPY, this also comes with a higher standard deviation: 29% for the two algorithms compared to 18% for SPY. Here, we once again see that using the average of the two algorithms results in better performance as the average lowers the standard deviation from 29% to 26%. Nevertheless, having a higher standard deviation than SPY shows that when the algorithms are able to obtain a higher annualized return, it is often taking more risks. To see whether it is profitable to take more risk, in that one can obtain a higher return relative to the "extra" risk one is exposed to, investors often look to the Sharpe Ratio.

	Algo1:GARCH	Algo2:Standard	Deviation	Average	SPY
Annualized Return	0.2110000		0.2134000	0.2235000	0.0815000
Annualized Std Dev	0.2938000		0.2948000	0.2609000	0.1836000
Annualized Sharpe (Rf=0%)	0.7183000		0.7238000	0.8567000	0.4439000
Worst Drawdown	0.4310669		0.5635507	0.4271594	0.5518943
Calmar Ratio	0.4894832		0.3786705	0.5232238	0.1476732
Ulcer Performance Index	1.2918619		0.9897342	1.4599161	0.6297990

Figure 12: Performance Measures

Given that both algorithms have similar annualized returns and similar standard deviation, they also have a similar Sharpe ratio at approximately 0.72, which is significantly higher that SPY's 0.44. However, as both algorithms have a Sharpe ratio below 1.0, they are neither particularly impressive as investors considers investments with a ratio below 1.0 to be sub-optimal (Maverick, 2020), despite higher ratios than SPY. Here we can once again notice that the performance can somewhat be improved by using the average, which gives a Sharpe ratio of 0.86.

With an annualized return of over 20%, the sub-optimal Sharpe ratio is most likely due to the algorithms being subjected to a relatively high degree of volatility. This is evident from the graph showing the strategies' drawdown, which shows that the first strategy loses approximately 43% while strategy 2 loses 56% of its value several times during the period from 2004 to 2019. For comparison, SPY lost 55% of its value during the Great Recession of 2008 – however, this is the only notable drawdown SPY had between 2004 and 2019. While using the average yields a less scary drawdown at 42.7%, the average still has several periods with a worse drawdown than SPY, as evident from Figure 12.

Up until now, the best performing strategy has seemed to be the average of the two strategies. The Calmar ratio confirms this, as the average has a ratio of 0.52 compared to 0.49 for the first strategy and 0.38 for the second. Furthermore, the average's Calmar ratio is higher than SPY's 0.15. The reason for this is because the Calmar ratio looks at the annualized rate of return compared to maximum drawdown, and we can see that the latter is significantly lower for the average, while the average also has a higher annualized return.

Because it uses a simple measure, maximum drawdown, the Calmar ratio is relatively simple to understand and use in a comparison of different algorithmic trading strategies. Its longer time frame also makes it less susceptible to short spikes of volatility, leading to more reliable results.

However, the focus on drawdown means that it only has a limited view of risk and may ignore general volatility, making it less statistically significant.

The final initial performance index we have used to compare the algorithm's and SPY is the Ulcer Performance Index. Like the Sharpe Ratio, this measure divides excess returns with a measure of volatility, but rather than using standard deviation, the performance index uses its own Ulcer Index. The advantage it thus enjoys over the Sharpe Ratio is that it focuses exclusively on downside risk related to the investment. Therefore, investors can more easily understand the volatility that really matters, or in other words, the volatility that may cause ulcers. The average once again yields the best performance, as it has the highest score of 1.46 compared to 1.29 by the first strategy, 0.99 by the second strategy, and 0.63 by SPY.

4.1.3 Transaction- and Management Costs

Finally, we can examine how the algorithms perform in terms of transaction and management costs. Figure 13 illustrates how the number of transactions for using the average of the two algorithms increases between 2004 and 2019. As we can see, although the number of transactions increases relatively evenly, there are periods where the curve is flat. This is where the algorithm holds a position in either VXZ or ZIV for an extended period, which presents a risk factor as these ETFs and ETNs, and especially ZIV, are highly volatile instruments.

In total, the average of the two algorithms uses 280 transactions, where half are trades in VXZ and the other half are trades in ZIV. This can be used to calculate the transaction costs. As ZIV was delisted on July 2nd, 2020, the closest estimate to ZIV's spread is SVXY's spread, which averages to 0.03% (etf.com, 2020). Meanwhile VXZ has a spread of 0.25% (etf.com, 2020). Therefore, the transaction costs would bring the algorithm's returns down by approximately 40%. In terms of management costs, VXZ maintains an expense ratio of 0.89% and SVXY's expense ratio is 1.38%, compared to SPY's 0.09% (etf.com, 2020). This implies that the algorithm would have approximately 1% higher yearly management costs than passively holding SPY. However, as the algorithm yields a 1880% cumulative return between 2004 and 2019, compared to SPY's 219%, these costs would not be able to bring down the returns to such an extent that using the algorithm would become less profitable than passively holding SPY.

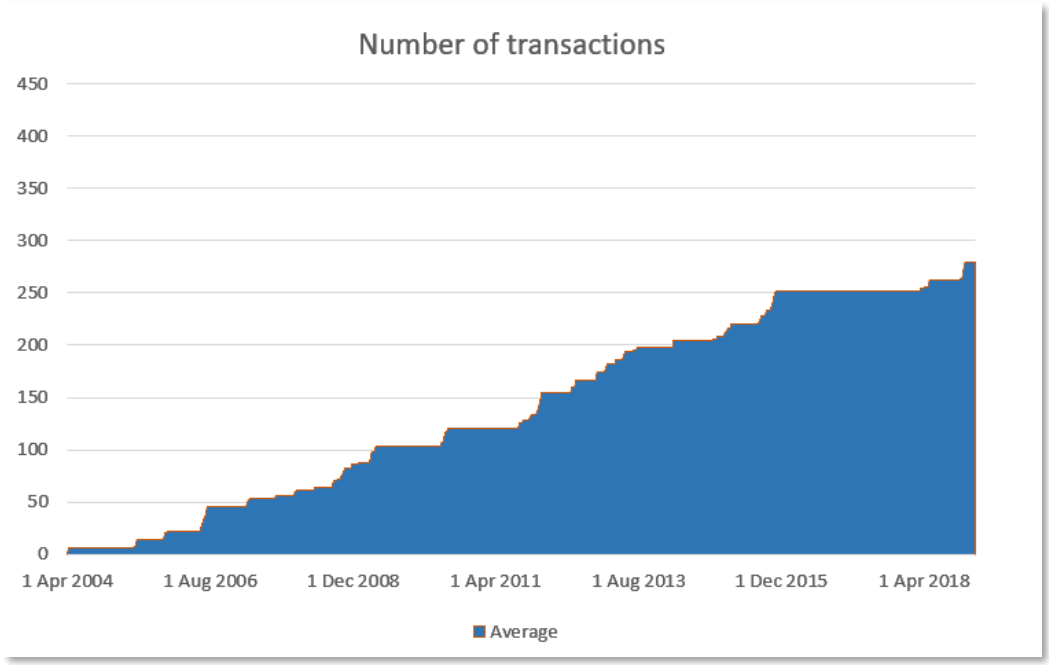


Figure 13: Number of Transactions - Average of models

4.2 Optimization of the Algorithm

In this section we will explore whether the average of the two strategies can be optimized for improved performance. We have decided to focus on reducing the risk.

4.2.1 Modifying the Algorithm

The main concern which should be addressed is the algorithm having prolonged periods of significant drawdowns. The latter often came because of holding a position in either VXZ or ZIV for an extended period. The implied risk can be visualised by changing ZIV in the algorithm for another inverse-VIX ETF – the SVXY. The results are visualised in Figure 14. After experiencing strong growth starting in 2016, SVXY fell by over 80% in after-hours on February 5th, 2018 (Stewart, 2018). As the algorithm would keep a position in SVXY during the uprun to this event, it made it impossible to exit the position during the after-hours crash, thus meaning you would lose most of your profits by using the algorithm with SVXY.

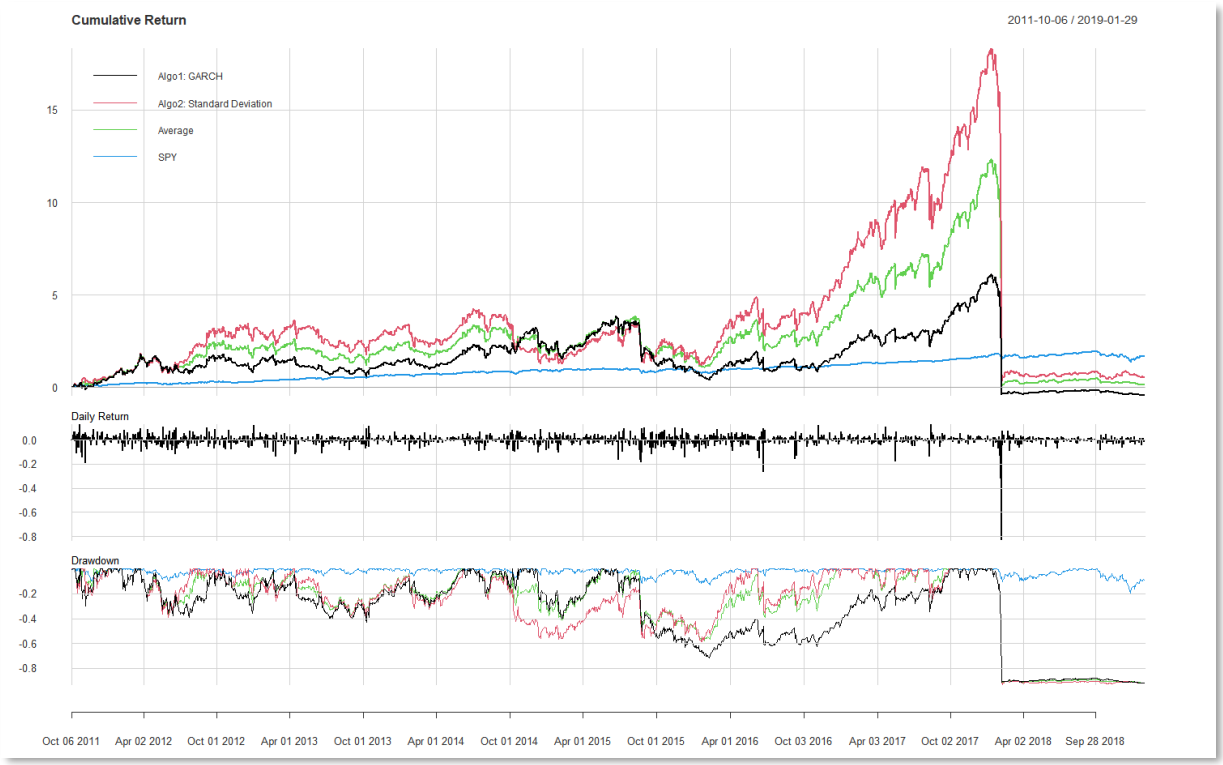


Figure 14: Cumulative Return – when applied to SVXY instead of ZIV

While ZIV did not experience the same fall in price as SVXY, it too is an inverse-VIX ETN, thus having a similar profile and therefore being exposed to similar risk. It must be noted that this risk is only targeting the inverse-VIX ETN, as these fall in value sharply during the occasional spikes which VIX has.

A possible solution to this is to gradually decrease the position held in the inverse-VIX ETN as the ETN's stock price goes up. An example of this is to decrease the held position by 1/3 when the inverse-VIX ETN goes up 20% in value since the position was entered. Then decrease the position by 1/3 again when the price has gone up 45%, and finally close the position when the price has gone up 85%. When applying this to SVXY, we obtain a more even chart, as seen in Figure 15.

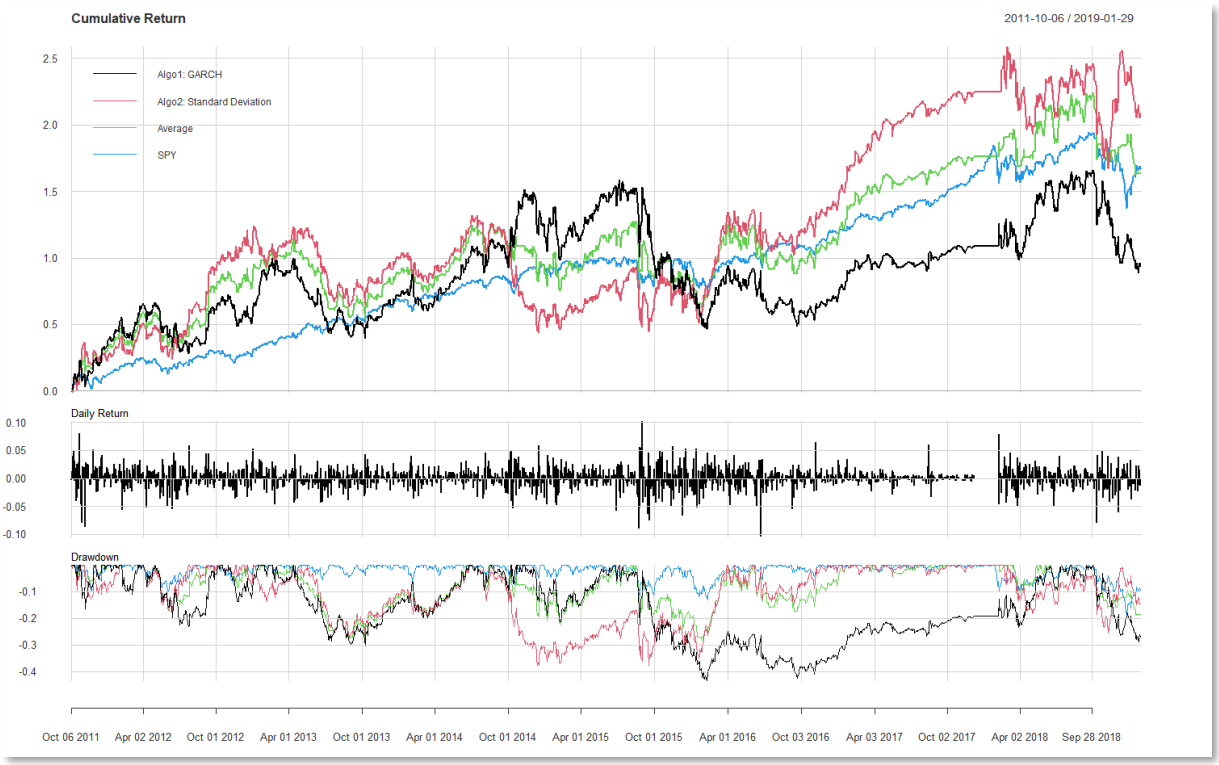


Figure 15: Cumulative Return – when applied to SVXY but adding the modification

With the modified algorithm, SVXY does not experience the same sharp fall in 2018, while it also yields significantly lower returns than previously. We can also observe that the algorithm with SVXY now follows a pattern close to that of SPY. However, as the algorithm with SVXY has now similar returns to SVXY while being more volatile, it is not a viable competitor to our previously developed algorithm with ZIV.

When applied to ZIV, the modified algorithm with continuous exits gives the following performance which can be seen in Figure 16.

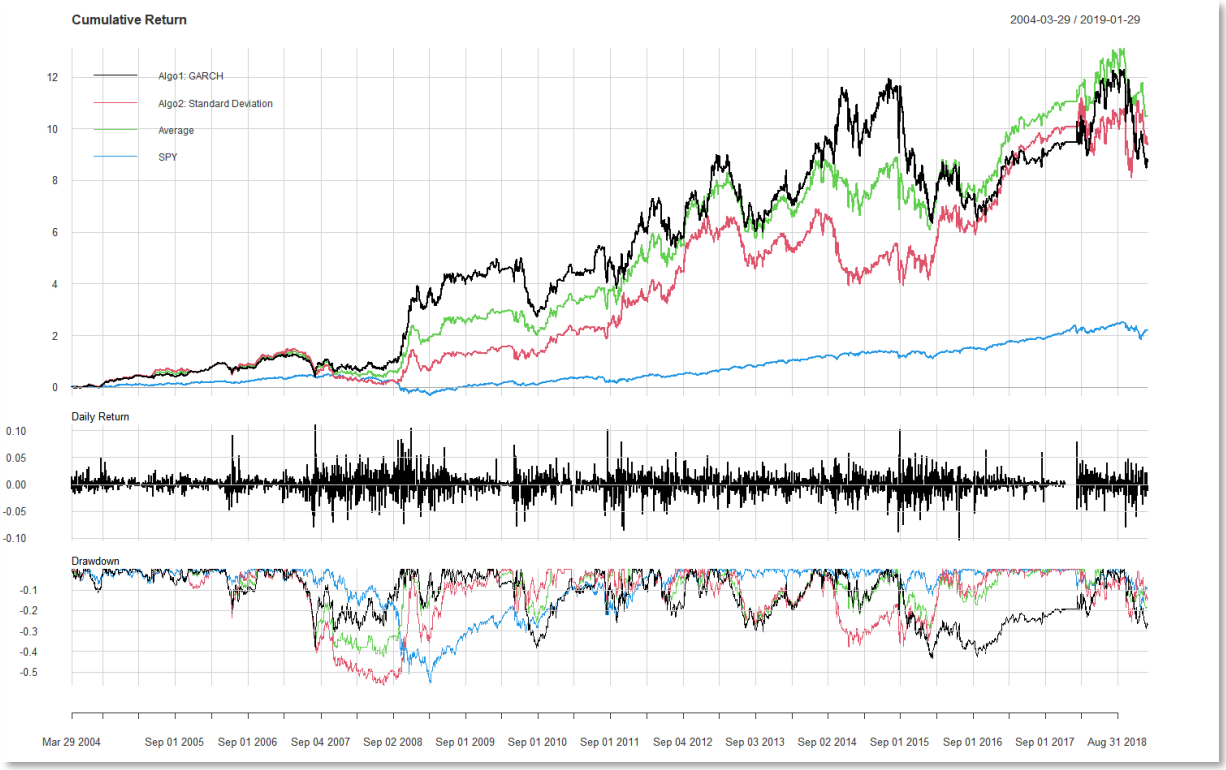
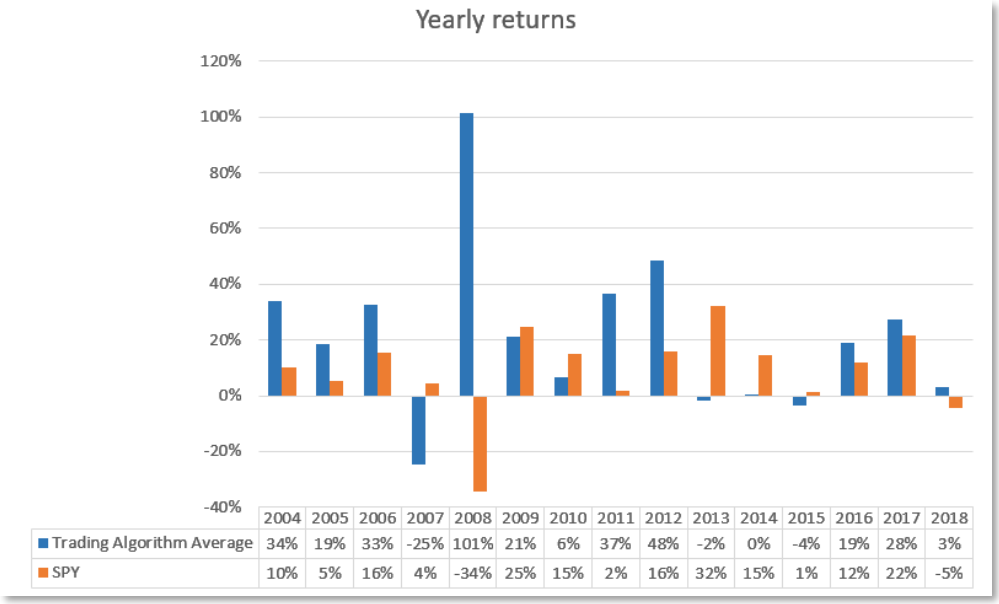


Figure 16: Cumulative Return – adding the modification to original algorithm with ZIV

From the figure we can see that although the profile when using ZIV has not changed drastically, the average of the two algorithms now experiences less volatility during the 2018 crash. This therefore minimizes the chances of the algorithm leading the user into a price runup ending with a crash.

4.2.2 Returns and Drawdowns

To further explore the modified algorithm, we can look at Figure 17 which again illustrates the year-by-year returns of the average of the two strategies plotted against the returns of SPY. Firstly, we notice that the overall returns of the average have decreased. However, we also see that the previous loss of -22% in 2018 is now turned into a positive 3% return. While the loss of -25% in 2007 remains, this loss was not based on a price runup ending with a crash, meaning it is difficult to improve the algorithm in this aspect. Overall, the modified algorithm yields more even returns and less risk of leading into a crash like that of 2018.



	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Trading Algorithm Average	34%	19%	33%	-25%	101%	21%	6%	37%	48%	-2%	0%	-4%	19%	28%	3%
SPY	10%	5%	16%	4%	-34%	25%	15%	2%	16%	32%	15%	1%	12%	22%	-5%

Figure 17: Yearly Return – Average of models after optimization

However, improvement presented by the modified algorithm comes at the cost of decreased returns. This is evident both in Figure 17 and Figure 18. The decreased returns are caused by the modified algorithm decreasing exposure when the held position in ZIV increases in value, which thus also limits the upside potential of the algorithm. To further explore this, we can look at Figure 18 which presents the performance measures of the average before and after modifying the algorithm, as well as the performance measures for SPY.

4.2.3 Sharpe, Calmar, and Ulcer

Figure 18 illustrates that while contributing to a decreased standard variation and reduced risk, the modified algorithm performs worse than the original algorithm regarding the Sharpe Ratio, Calmar Ratio, and the Ulcer Performance index. This implies that the decreased risk caused by the modified algorithm is disproportional with the decreased returns. Nevertheless, the modified algorithm keeps largely outperforming SPY in all three of these performance measures, and the decrease should therefore be acceptable as it reduces the systematic risk of the algorithm leading the user into an ETF or ETN crash. Additionally, it can also be noted that the modified algorithm has an unchanged worst drawdown, which is due to the loss in 2007 being unchanged – which is where the worst drawdown occurs.

	Average (original)	Average (continuous exits)	SPY
Annualized Return	0.2235000	0.1790000	0.0815000
Annualized Std Dev	0.2609000	0.2409000	0.1836000
Annualized Sharpe (Rf=0%)	0.8567000	0.7429000	0.4439000
Worst Drawdown	0.4271594	0.4271594	0.5518943
Calmar Ratio	0.5232238	0.4190473	0.1476732
Ulcer Performance Index	1.4599161	1.2925295	0.6297990

Figure 18: Performance Measures - After optimization

Finally, we can look at how the modified algorithm affects the number of transactions. Figure 19 shows the number of transactions for the original average and after the modification has been implemented. We can see that the number of additional transactions caused by the modification increases steadily over time until the period between 2016 and 2018 when this number starts increasing more rapidly. During this period, the original algorithm holds a position in ZIV for an extended period, leading through the price increase and into the volatility crash in 2018. The long position in ZIV can be seen by the flat blue curve in this time frame. Meanwhile, when forcing the algorithm to reduce the position as its value increases, we see that the orange curve increases, thus adding transactions and exiting the position before the crash.

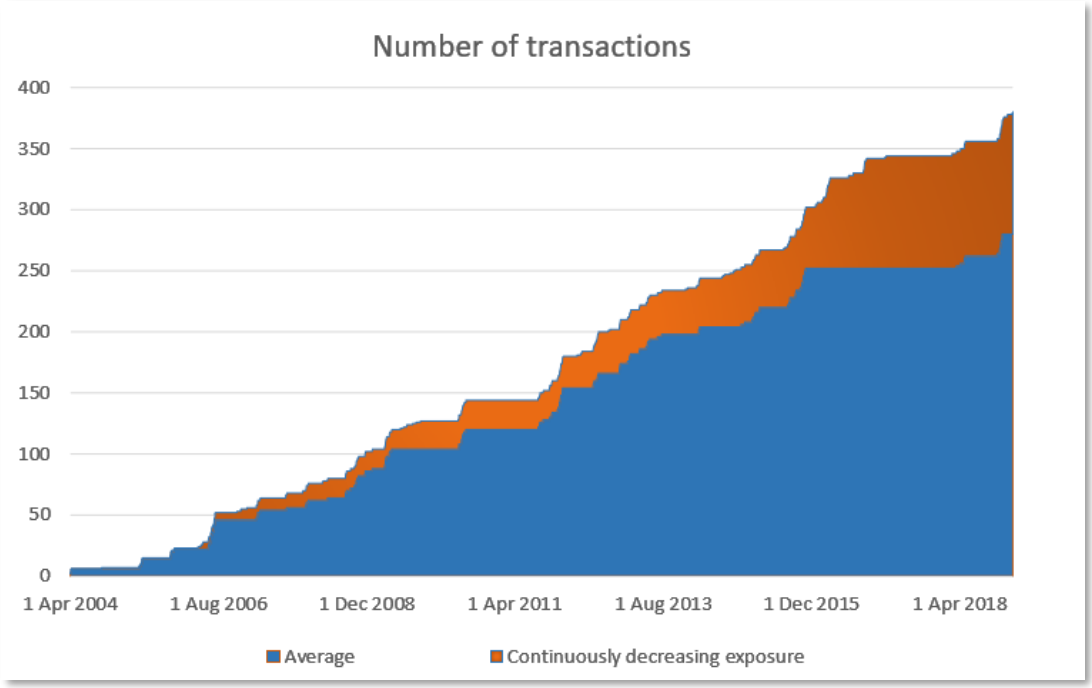


Figure 19: Number of Transactions - After optimization

4.2.4 Transaction- and Management Costs

In terms of transaction costs, the modified algorithm completes an additional 100 trades between 2004 and 2019. Keeping in mind the 280 transactions completed by the original algorithm, implementing this modification would increase the original transaction costs by approximately 35%. The management costs would mostly be unaffected, as the only prolonged time frame when the modified algorithm fully exits a position without entering a new one is just before 2018. In total however, the increased transaction costs would not make the algorithm less profitable than passively holding SPY, as the modified algorithm produced cumulative returns of 1047%, very comfortably beating SPY's 219%. This shows that by using a modified version of the average of the two algorithms, we can achieve returns greatly exceeding a passive hold strategy, while also minimizing the risk of the algorithm walking into a crash.

5 Conclusion

The purpose of this paper has been to explore the field of algorithmic trading and the process of designing a trading algorithm. In the beginning, we introduced the theoretical framework for designing such an algorithm, as well as the measures we would use in the analysis. Furthermore, we presented the two Volatility Risk Premium strategies which we would use as an example for a trading algorithm, as well as the implementation of these strategies, which was developed by Ilya Kipnis.

In the Performance Analysis, we found that both strategies outperform a buy and hold strategy for SPY based on the measures we used – the only exception being the two strategies requiring a higher level of risk. Further, the average of the two algo-strategies outperformed using any of them individually. For the Optimization part, we therefore focused on the average of the two strategies. Our desire here was to reduce the implied risk of using the algorithm, especially in regard to the volatility crash in 2018. By incorporating a concept of continuously decreasing the position in the algorithm when the inverse-VIX, ZIV ETN increased in value, we managed to lower the standard deviation and risk, and prevent the algorithm from walking into the volatility crash of 2018. However, this improvement came at the cost of reduced returns, worsening the other performance measures we used. That being said, the optimized strategy still outperformed the SPY, S&P500 ETF, by producing a nearly five times greater return, and now with reduced risk.

In terms of weaknesses concerning our analysis, transaction costs are worth mentioning. Transaction costs are highly complex and consists of more than our analysis has taken into consideration. We have mainly focused on how spreads can affect performance, but other measures such as cost of impact and time cost can also make up a significant portion of the total transaction costs. The forementioned costs will be a function of the size of the fund; for small sums of capital, they will play a minor role.

Finally, while it must be kept in mind that past performance is no guarantee of future results, with the average of the two strategies massively outperforming the SPY over a 15-year period, it suggests that the overall algorithm has potential for being a profitable one, and one can suggest that it could continue producing positive returns.

References

Bloomenthal, A. (2020, March 20). *Investopedia*. Retrieved from World's Top 10 Hedge Fund Firms: https://www.investopedia.com/articles/personal-finance/011515/worlds-top-10-hedge-fund-firms.asp

Chen, J. (2020, March 16). *Exchange Traded Funds - ETFs*. Retrieved from Investopedia: https://www.investopedia.com/terms/e/etf.asp

Chen, J. (2020, March 24). *Exchange Traded Notes - ETN*. Retrieved from Investopedia: https://www.investopedia.com/terms/e/etn.asp

Chen, J. (2020, October 6). *Treasury Bond (T-Bond)*. Retrieved from Investopedia: https://www.investopedia.com/terms/t/treasurybond.asp

Cooper, T. (2013). Easy Volatility Investing. *SSRN Electronic Journal*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2255327

Credit Suisse. (2020, December 11). *ZIV*. Retrieved from Velocityshares: https://www.velocityshares.com/etns/product/ziv/

etf.com. (2020, December 10). *ETF.com*. Retrieved from VXZ iPath Series B S&P 500 VIX Mid-Term Futures ETN: https://www.etf.com/VXZ

etf.com. (2020, December 10). *ETF.com*. Retrieved from SVXY ProShares Short VIX Short-Term Futures ETF: https://www.etf.com/SVXY

etf.com. (2020, December 10). *ETF.com*. Retrieved from SPY SPDR S&P 500 ETF Trust: https://www.etf.com/SPY

ETFGI. (2019, July 15). *ETFGI*. Retrieved from ETFGI reports assets invested in the Global ETF industry: https://etfgi.com/news/press-releases/2019/07/etfgi-reports-assets-invested-global-etf-and-etp-industry-reached

Goldman Sachs. (2018). *Top of Mind: Liquidity, Volatility, Fragility*. New York: The Global Investment Research Division of Goldman Sachs. Retrieved from https://www.gsam.com/content/dam/gsam/pdfs/sg/en/commentary/GS_Top%20of%20Mind_June.pdf?sa=n&rd=n

Hargrave, M. (2020, April 25). *Sharpe Ratio*. Retrieved from Investopedia: https://www.investopedia.com/terms/s/sharperatio.asp

Harris, L. (2003). *Trading and Exchanges: Market Microstructure for Practitioners.* New York: Oxford University Press.

Kawa, L. (2019, December 12). *Ghost of XIV Rises: A New Short Vol Product Is In The Pipeline*. Retrieved from Bloomberg: https://www.bloomberg.com/news/articles/2019-12-12/ghost-of-xiv-rises-a-new-short-vol-product-is-in-the-pipeline

Kenton, W. (2020, July 1). *The Calmar Ratio*. Retrieved from Investopedia: https://www.investopedia.com/terms/c/calmarratio.asp

Kipnis, I. (2018, December 3). *GARCH and a rudimentary application to Vol Trading*. Retrieved from QuantStrat TradeR: https://quantstrattrader.com/2018/12/03/garch-and-a-rudimentary-application-to-vol-trading/

Maverick, J. (2020, October 16). *What is a Good Sharpe Ratio?* Retrieved from Investopedia: https://www.investopedia.com/ask/answers/010815/what-good-sharpe-ratio.asp

Patterson, S. (2011). *The Quants: How a new breed of math whizzes conquered Wall Street and nearly destroyed it.* New York: Crown Business.

Quantstart. (2020, December 12). *Autoregressive Moving Average ARMA(p, q) Models for Time Series Analysis - Part 1*. Retrieved from Quantstart: https://www.quantstart.com/articles/Autoregressive-Moving-Average-ARMA-p-q-Models-for-Time-Series-Analysis-Part-1/

Remsburg, J. (2020, June 2). *Investorplace*. Retrieved from The Power of a Numbers-Based System: https://investorplace.com/2020/06/the-power-of-a-numbers-based-system/

Samuelsson, H. (2020, November 17). *10 Pros and Cons of Algo Trading*. Retrieved from The Robust Trader: https://therobusttrader.com/pros-and-cons-of-algo-trading/?fbclid=IwAR3yEk3Z4Ww7ZcB2DKJ4axiJDEdund8iSoHSs0TjkDGuh37usHy35N3yfDw

SEC & CFTC. (2010). *Findings regarding the market events of may 6, 2010.* Washington, D.C.: Commodity Futures Trading Commissioin & Securities and Exchange Commission. Retrieved from https://www.sec.gov/news/studies/2010/marketevents-report.pdf

SEC. (2020). *Staff Report on Algorithmic Trading in U.S. Capital Markets.* Washington: U.S. Securities and Exchange Commission. Retrieved from https://www.sec.gov/files/Algo_Trading_Report_2020.pdf

Seth, S. (2020, March 7). *Basics of Algorithmic Trading: Concepts and Examples*. Retrieved from Investopedia: https://www.investopedia.com/articles/active-trading/101014/basics-algorithmic-trading-concepts-and-examples.asp

Smartereum. (2019, June 26). *The Quest to Bring Crypto Automated Trading to the Masses*. Retrieved from Smartereum: https://smartereum.com/58578/the-quest-to-bring-crypto-automated-trading-to-the-masses/

Smith, T. (2020, March 10). *Autocorrelation*. Retrieved from Investopedia: https://www.investopedia.com/terms/a/autocorrelation.asp

Stewart, E. (2018, February 27). *The next big bad vet might be lurking on Wall Street - and we just got a glimpse of it*. Retrieved from Vox: https://www.vox.com/business-and-finance/2018/2/27/17014082/market-crash-inverse-volatility-vix-xiv-svxy

Zuckerman, G. (2019, November 5). *Financial News London*. Retrieved from The making of Jim Simons: the world's greatest investor: https://www.fnlondon.com/articles/the-making-of-jim-simons-the-worlds-greatest-investor-20191105