Charles University

Faculty of Social Sciences
Institute of Economic Studies

MASTER'S THESIS

Trading strategies based on estimates of conditional distribution of stock returns

Author: **Bc. Adam Sedlačík**

Supervisor: **doc. PhDr. Jozef Baruník, Ph.D.**

Academic Year: **2017/2018**

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain a different or the same degree.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, July 27, 2018

Signature

Acknowledgments

I am especially thankful to my supervisor Jozef Baruník. He answered every question practically immediately and always provided some insightful notes which pushed me further.

I thank my parents not only for their financial support, which largely helped me to stay calm a focus on my studies but especially for their psychological support - the most valuable help they could give me.

I am also thankful to my girlfriend Michaela, who had to cope with all my general brainstorming questions, which helped me to look at things differently.

Abstract

In this thesis, a new trading strategy is proposed. By the help of quantile regression, the conditional distribution functions of stock market returns are estimated. Based on the knowledge of the distribution the strategy produced buying and selling signals which together with a weight function derived from exponential moving averages determines how much and when to buy or sell. The strategy performs better than the market in terms of absolute return and the Sharpe ratio in-sample, but it does not provide satisfactory results out-of-sample.

JEL Classification	G11, G14, G17
Keywords	quantile regression, trading strategy, conditional distribution function estimation, GARCH

Author's e-mail	adam.sedlacik@yahoo.com
Supervisor's e-mail	barunik@fsv.cuni.cz

Abstrakt

Tato práce vytváří novou obchodní strategii. Pomocí kvantilových regresí jsou odhadnuty podmíněné distribuční funkce tržních návratností. Na základě odhadnutých distribucí jsou produkovány signály pro nákup a prodej, které společně s funkcí vah, jež jsou odvozené od exponenciálních klouzavých průměrů, určují množství a čas nákupu a prodeje. Strategie má lepší výkonnost než trh na datech, na kterých byla odhadována, avšak při out-of-sample testování neposkytuje uspokojivé výsledky.

Klasifikace JEL	G11, G14, G17
Klíčová slova	kvantilové regrese, obchodní strategie, odhady podmíněné distribuční funkce, GARCH

E-mail autora	adam.sedlacik@yahoo.com
E-mail vedoucího práce	barunik@fsv.cuni.cz

Contents

List of Tables

List of Figures

Acronyms

ACF Auto-Correlation Function

AIC Akaike information criterion

ARMA Auto-regressive Moving Average

BW Buying Weight

EMA Exponential Moving Average

GARCH Generalized Autoregressive Conditional Heteroskedasticity

OLS Ordinary Least Squares

PACF Partial Auto-Correlation Function

SW Selling Weight

VaR Value at Risk

Master's Thesis Proposal

Author	Bc. Adam Sedlačík
Supervisor	doc. PhDr. Jozef Baruník, Ph.D.
Proposed topic	Trading strategies based on estimates of conditional distribution of stock returns

Motivation

In order to effectively invest or trade on financial markets it is needed to have a well-defined model and to stick to it while making investment decisions without an emotional impact. The model needs to be sufficiently simple but also include relevant available data in order to provide good results when it comes to real investment.

The majority of trading models in current literature estimate the expected stock price and its variance as a measure of risk. It is completely correct as long as the returns are normally distributed which is, however, one of the restrictive assumptions which are hardly met in reality. The distributions of returns have often higher kurtosis and heavy tails causing increase of probability of extreme losses. That was the reason why researchers and also practitioners diverted from the classical OLS model and developed more advanced techniques. They relaxed possibly all assumptions and created nonparametric models which take into account only the information contained in the data.

The best solution would be of course to know the future excess return but as this will most probably be never possible the next best solution is to know the distribution of returns. The knowledge of the distribution helps the investor assess the risk and behave accordingly. The goal of our thesis is first to predict conditional distribution function of excess returns based on various macroeconomic data, company accounting numbers and basic trading measures and second, based on the predictions, to develop a trading strategy which could systematically outperform the market returns - e.g. buy and hold strategy.

Hypotheses

Hypothesis #1: The new trading strategy, based on estimates of future distribution of excess returns, systematically outperforms the market in very long run.

Hypothesis #2: The newly proposed trading strategy provides statistically better results than usual trading methods.

Hypothesis #3: : Self-learning trading model provides better results than empirical or conceptual model.

Methodology

The model used will be similar to the one used in Foresi and Peracchi (1995), particularly the logistic additive regression model. We will use a logit additive regression where instead of particular parameters s smooth functions will be estimated. For this approach kernel or spline regression can serve as a good tool. Instead of for expected value this regression will be repeated for sufficiently large number of quantiles. Each regression will estimate the probability of future excess return being below or above particular, previously defined, value. All these regressions together provide an estimate of distribution function of excess returns conditional on preselected variables. Based on the estimate of conditional distribution function of excess returns I will find out whether it is meaningful to use it for a development of trading strategy and whether this strategy outperforms at least buy and hold strategy. As other studies have shown (Foresi and Peracchi 1995; Thomas Q. Pedersen 2015) different quantiles of distribution are affected unequally by selected predictors thus opening space for more accurate trading strategy.

The analysis will be conducted mainly on daily returns of quoted companies and it will be also accounted for transaction cost associated with purchasing and selling of stocks.

As a possible extension we may try to create a model which is capable of learning itself, model which incorporates its past mistakes and improves the future estimates. For this purpose we try to incorporate existing learning algorithms.

Expected Contribution

We expect to develop new trading strategy based on distribution estimation and possibly learning algorithms. The goal is to perform out-of-sample tests and conclude whether it is first possible to estimate future distribution function of returns to a level sufficient for outperforming other trading strategies and second create self-learning algorithm useful for everyday investing.

Outline

1 Introduction

2 Motivation

2.1 Trading generally and strategies

3 Theoretical background of used methods and models.

3.1 General additive models - logistic additive model

3.2 Non-parametric estimation

3.3 Self-learning algorithms

4 Data description

5 Empirical analysis

6 Conclusion

Core bibliography

Foresi, Silverio, and Franco Peracchi. The conditional distribution of excess returns: An empirical analysis. Journal of the American Statistical Association 90, no. 430 (1995): 451-466.

Friedman, Jerome, Trevor Hastie, and Robert Tibshirani. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). The annals of statistics 28, no. 2 (2000): 337-407.

Hastie, Trevor J., and Robert J. Tibshirani. Generalized additive models. Vol. 43. CRC press, 1990.

Friedman, Jerome, Trevor Hastie, and Robert Tibshirani. The elements of statistical learning. Vol. 1. Springer, Berlin: Springer series in statistics, 2001.

Gallant, A. Ronald, Peter E. Rossi, and George Tauchen. Stock prices and volume. Review of Financial studies 5, no. 2 (1992): 199-242.

Fama, Eugene F., and G. William Schwert. Asset returns and inflation. Journal of financial economics 5, no. 2 (1977): 115-146.

JW Yang, J Parwada. Predicting stock price movements: an ordered probit analysis on the Australian Securities Exchange. Taylor & Francis Journals, vol. 12, no. 5 (2012): 791-804.

Thomas Q. Pedersen. Predictable Return Distributions. Journal of Forecasting 34, no. 2 (2015): 114-132.

Author

Supervisor

Chapter 1

Introduction

In the world the business activity is revolving around stock markets, investing in securities and generally allocating money to places where it is needed. Stock markets serve as a tool for the allocation and provide necessary liquidity for market participants (Bernstein 1987). Liquidity, as a measure of ease to exchange assets into money, is a vital means for proper functioning of every stock exchange. Investors as well as traders are an integral part of it.

Every investor or trader seeks an individual goal on the stock market. There may be a great number of reasons. The investor may want to invest with an objective of long-term value preservation or the purpose may be the desire to help the companies grow - to offer them capital which they can use in their operations. By investment, the investor may protect their wealth. They can invest the excess resources to postpone the consumption. The satisfaction of being active on the stock market may also carry some value. Even gambling can be an objective. All of the previously mentioned purposes possess a common feature. The investor or trader offers their excess money in exchange for some expected risk premium. The size of the risk premium is usually dependent on the type of investment instrument. Stocks bear a higher risk than bonds simply because the shareholders claim the residual income - what has left after all obligations are met.

Investors usually allocate their resources with a predefined rule or strategy. The rules tell them what it could be expected and offer the ability to assess their investment strategy performance. The investment strategies include immense amount of possibilities how to invest. Generally, the investor wants to maximize profit, minimize loss or reach some consensus of both. This may be done by portfolio selection strategies or with diversification which reduces the unsystematic risk. However, there still remains some form of systematic risk which cannot be diversified.

The objective of this thesis is to develop a trading strategy which offers the investor better returns than the market or better risk-weighted return thus mitigating the systematic risk or getting better returns with the same level of risk. By the help of quantile regression estimates the thesis tries to sufficiently precisely forecast the distribution of market returns and based on the estimates to develop the desired trading rule.

The thesis is organized as follows. Chapter 2 covers the literature review and current findings. In chapter 3 the data are described. Chapter 4 provides the complete methodology used in this thesis - the methods of conditional distribution function estimation and the setup of the newly proposed trading strategy. The results of the model are reviewed in chapter 5 and the chapter 6 interprets the results in the previous chapter and motivates the reader for further research. Finally, the chapter 7 summarizes the proposed strategy and concludes the thesis.

Chapter 2

Literature review

There is a rather large amount of papers and publications trying to invent a way how to predict future stock returns (Rapach & Zhou (2013), Cenesizoglu & Timmermann (2007), Cenesizoglu & Timmermann (2008), Yang & Parwada (2012)). The usual one, which is still used in practice, is to estimate the first two moments of a distribution of returns series, i.e. the expected value of returns and its standard deviation as a measure of risk. However, this procedure largely limits the model ability to reflect the reality since it has to expect that the process comes from a predefined distribution, particularly the normal Gaussian. In any case, some investors are not only interested in the first two moments but also in the third (skewness) and fourth (kurtosis). It has been shown, for example, that these moments have significant effects on cross-sectional variation in US stock returns and its expected value (Harvey & Siddique (2000), Dittmar (2002)). However, one still needs to assume the data come from a specific distribution, thus, limiting the model. Consequently, new studies have concluded that imposing the shape of the distribution is not a good solution, in other words, it proves that returns are not predictable with the help of this simple model specifications.

During the years statisticians have developed more advanced methods of predicting distributions trying to relax one or more of the implied assumptions. They devised quantile regressions which estimate the effects of explanatory variables separately for different quantiles (Koenker & Bassett Jr (1978), Koenker & Hallock (2001)). This avoids the aggregation of data into a prespecified distribution shape and allows for more accurate predictions for each quantile. Classical OLS minimizes the sum of squares of the difference between the realized values of the dependent variable and the fitted values predicted by the linear model. OLS is quite sensitive to outlier observations thus it is possible to perform a median regression which instead of minimizing the sum of squares minimizes the sum of absolute distances. Quantile regression goes further and with the help of weights estimates various quantiles of the response variable. This procedure is especially helpful in situations when the conditional distribution of the response variable does not have any specific shape.

In general, most of the studies which use quantile regression for estimation come up with a solution that some market or economy variables have an effect on the left or right tail of the distribution of returns but have no statistically significant effect in the center of the distribution. This is of practical use since the usual goal is to outperform the market in a very long run - trying to outperform the bull market is a thing which lacks punchline since the investor is satisfied that he earns a profit. More importantly, the investor has to know or at least significantly predict the start of a bear market. The benefits here are relatively straightforward.

There are two broad types of investment-management strategies: active and passive management. A passive investment strategy is based on buying various types of assets to build a well-diversified portfolio without further and deep analysis of individual assets. A classical example is the buy-and-hold strategy using the naive 1/N portfolio. On the other hand, the active management aims to scrutinize each asset with the help of possibly all related public information. Of course only to such a level at which the costs of gathering and implementing new information are lower than the expected benefits.

Finally, information is needed to make a proper investment decision. Searching through a number of studies has shown what could be significant in predicting stock market returns. The potential candidates are trade size, trade indicator, bid-ask spread, volume, the depth of the market, the price of gold, silver or oil, various exchange rates, mainly EUR/USD, PE ratio, PS ratio, dividend-price ratio, inflation, unemployment etc. Still, just little is known about which part of the distribution is predictable by which states (Cenesizoglu & Timmermann 2008). The whole topic needs a rougher analysis.

A lot of studies have shown that stock market prices (or log prices) witness unit root. This conclusion, however, does not mean that the prices follow a random walk. To follow a random walk the series has to be uncorrelated or serially independent - there can be dependence in higher moments. To find this out a nonparametric approach may be a good solution.

There are also studies which tried to forecast the excess returns using quantile regression or other advanced techniques which do not restrict the researcher to a parametric model. We have to bear in mind that there is still an extensive amount of literature which concludes that a number of economic variables do not have an effect on market returns (Welch & Goyal 2008). However, these studies do not use quantile regression for its purpose. Pedersen (2015) estimated the return distribution of stocks and bonds using 8 economic state variables among which were the dividend-price ratio, inflation, short-term interest rate (3 months treasury bills) or the term spread using quantile regression. He concluded that the economic state variables forecast the stock return in different ways in terms of location shifts, volatility and skewness, next that most economic state variables forecast the stock return distribution quite well out-of-sample, while for bonds the forecast performance is less good and that the forecasting performance of various economic state variables varies across the return distribution. He also presented that the density forecasts based on quantile regression, in general, show better performance in comparison with forecast assuming the normal distribution. The most important point is the rejection of independence of stock returns showing that there may be a space for profitable trading.

According to Pedersen (2015) it can happen that a variable is a good predictor of, for example, the left tail of the distribution but does not have a significant effect along the middle part or right tail of the distribution. He also concluded that forecasts based on quantile regression, in general, show better performance than density forecasts based on a normal distribution with forecasted mean and variance. Pedersen (2015) together with Koenker & d'Orey (1987) confirmed the "flight-to-quality" pattern which means that investors tend to reduce their stock holdings in periods of higher market uncertainty and invest more in bonds. The stock market uncertainty may be proxied by the return variance of S&P 500 Index price movements.

Yang & Parwada (2012) used an ordered-probit-GARCH model and concluded that it is possible to predict in 71 % cases the direction of price movements. However, the key point here is whether the upward and downward movements have not a statistically significant difference in magnitude since this is important for the building of a profitable trading strategy. An important part of the trading strategy could be also the volume and number of trades. Yang & Parwada (2012), for example, tested the exogeneity of the duration of time between trades using an Auto-regressive Conditional Duration specification and concluded that it is negatively related to the price change. Also, the duration between two consecutive trades is not exogenous but depends on other market attributes (Easley & O'hara 1992). Current studies using high-frequency data try to approximate the price movements to such a detail that it could be considered continuous. However, Yang & Parwada (2012) have shown that the prices do not move continuously in reality thus a continuous approximation is not needed. Another influential variable is the trade imbalance representing the proportion of buy and sell limit orders placed on the market which has a strong positive correlation with the conditional price changes - in other words, it confirms the demand/supply relation. Generally, Yang & Parwada (2012) concluded that the price moves are predictable but as previously mentioned their findings are qualitative but not quantitative. They only looked at the direction of price change but not for magnitude.

Easley & O'Hara (1987) created a theoretical framework which states that there is a positive relationship between trade size and returns. This finding is verified by an empirical studies made by Hasbrouck (1991) or Chordia & Shivakumar (2002). Also, a quantity of trade influences the price movements (Easley & O'Hara 1987). The simple explanation states that large trades are usually correlated with private information about a stock.

Tests based on autocorrelation are not effective against nonlinear processes with zero autocorrelation (Granger & Andersen 1978).

The whole topic about predictability revolves around the efficiency of the market. There are three forms of efficiency: weak-form, semi-strong-form, and strong efficiency. In the weak-form efficiency, the future prices cannot be predicted using prices from the past resulting in a uselessness of technical analysis. The semi-strong efficiency states that all available public information is useless for predicting the future price. Finally, when the markets are strongly efficient even the insider information is of no use. Generally, studies concerning the market efficiency conclude that the markets are weakly efficient but are not strongly efficient (Cochrane 2008), some have also found out that even the available public information represented by different valuation ratios is of no use and the out-of-sample performance is insignificant (Welch & Goyal 2008). On the other side, there are also studies which inferred that the stock market returns may be predictable using only the historical prices. The reasons for that are temporary market inefficiencies (Park & Irwin 2007) or temporary forecastability (Timmermann & Granger 2004). Positive forecasting results of the technical analysis may be achieved during certain time periods.

An issue apart from the estimation of conditional distribution is the estimation of the joint distribution of explanatory variables. Amisano & Giacomini (2007) used a weighted likelihood ratio test but the test is still not sufficient and there is a need for an approach that can take the dependence structure into account when providing an estimate of the joint distribution.

Beside all that the rigid assumption can be also avoided by ARCH or GARCH models (Bollerslev *et al.* 1992). These models are a good solution for conditional stock market volatility estimation. Or to overcome practically all assumptions and even function specifications some researchers try to predict the future market returns using neural networks, particularly the feedforward neural networks. These methods received a lot of attention in recent years but the application of them goes back a couple of decades (Azoff (1994), Trippi & DeSieno (1992)). The indisputable advantage of neural network systems is that it can basically simulate any function which could potentially predict the stock market movements thus offering opportunities for profitable trading (Hornik 1993).

The strategy which will be developed in this thesis may be considered a momentum strategy. The well-known adage: "buy low, sell high" is basically the principal of it. It could be extended by the philosophy of Richard Driehaus: "buy high, sell higher". The idea of it states that good performing stock will also perform well in the short-term period. For example, Jegadeesh & Titman (1993) have found out that buying stocks which performed poorly and selling stocks which performed well in past 3 to 12 months is a profitable strategy. It was first considered as market inefficiency because it enabled to create an idea that the market returns are predictable. However, several explanations occurred during the years - Conrad & Kaul (1998) have argued that the presumed predictability or the momentum returns are just a matter of cross-sectional differences in expected returns. On the other hand, the functionality of momentum strategies may have a psychological effect. Kahneman & Tversky (1982) made a behavioral statement that people tend to overreact to available information thus a company which price increased in the past will probably grow in the short future.

Taking into account any strategy it is needed to properly determine its profitability. This means to fully assess the performance it is needed to find the drivers of that performance and explain them. Without a proper reasoning, the analyst may reach a conclusion that the strategy performs well even though it was produced by pure data snooping bias.

The unpredictability of stock market returns is usually perceived as a stock market efficiency (Fama 1991). This, however, does not have to be true. It may only signify that the proper model for testing the efficiency has not been found and together with that - a functional long-term trading strategy has not been developed.

Chapter 3

Data

The first idea already mentioned in the proposal was to create a trading strategy based on various macroeconomic, fundamental and other market data. The author wanted to develop a trading strategy which takes into account inflation, unemployment, GDP, PE or PS ratios, dividends etc. However, there arises a problem with the data frequency. A great amount of macroeconomic data series, which were originally intended to use, are not available at daily frequencies. Some of them are observed weekly but the majority is published only each month. For example, the M2 Money Stock containing financial assets held principally by households is disclosed weekly, the Civilian Unemployment Rate is a monthly series, the same holds for the Consumer Price Index as well as Real Personal Consumption Expenditures. The listing of macroeconomic series at monthly frequencies could continue. And, besides that, there is the GDP, which is a quarterly series. To all this, each series is also published in different time. Some of them lag 5 trading days (Civilian Unemployment Rate, M2 Money Stock), some two weeks (Consumer Price Index) and some a whole month (Real Personal Consumption Expenditures). In order to include them in the analysis, one would not need the first lag but the second to make a prediction. There also arises a problem how to combine the series together - how to set the individual observation "rows" and whether they do not influence each other after the data of the first series are published.

There are solutions of how to increase the frequency of the data or how to fill the missing observations. Nonetheless, to make a daily series out of monthly data it would require to create for 1 realized observation additional approximately 21 artificial data points. This could be done by any type of interpolation e.g. cubic splines but it would hardly meet the reality because it would be a very rough approximation.

At this point arises a question why not to use monthly data and develop the strategy on lower frequencies. This is certainly a possibility because most of the series are available for a couple of decades which offers a sufficient amount of data to conduct an analysis. Taking an example, the M2 Money Stock is available from 1980 until today. This constitutes 37 years which totals 444 monthly observations. It is undoubtedly enough for Ordinary Least Squares or for other mean regression but is not a sufficient number of observations if one would like to estimate the shape of the distribution. For example, if the analyst wanted to estimate the 5 % lower quantile then they would have only 22 observations in this region to model that particular part of the return distribution. If they wanted to estimate even the 1 % quantile, they would have only 4,4 observations for that region which is definitely not enough. The same problem arises with the fundamental data of individual companies.

Because of the poor data availability together with the focus of this thesis, it was decided to use data which posses at least daily frequency. This decision considerably reduced the set of possible series which could be selected. Due to the lag of any fundamental data in daily frequency, the whole model will be developed and tested on the S&P 500 innovations[1]. The other variables used for modeling and available at daily frequencies are 10-Year Breakeven Inflation Rate (FRED economic data code: T10YIE), 10-Year Treasury Constant Maturity Minus 3-Month Treasury Constant Maturity[2] (FRED code: T10Y3M) and the volatility index VIX. "The breakeven inflation rate represents a measure of expected inflation derived from 10-Year Treasury Constant Maturity Securities and 10-Year Treasury Inflation-Indexed Constant Maturity Securities. The latest value implies what market participants expect inflation to be in the next 10 years, on average"(Federal Reserve Bank of St Louis 2018). The definition of 10-Year Treasury Constant Maturity Minus 3-Month Treasury Constant Maturity is not needed. "The VIX Index is a calculation designed to produce a measure of constant, 30-day expected volatility of the U.S. stock market, derived from real-time, mid-quote prices of S&P 500® Index (SPX℠) call and put options. On a global basis, it is one of the most recognized measures of volatility – widely reported by financial media and closely followed by a variety of market participants as a daily market indicator."(CBOE VIX 2018).

[1] S&P 500 index is also labeled as GPSC, or it could be referenced to as "the market".

[2] The series is also called the term spread.

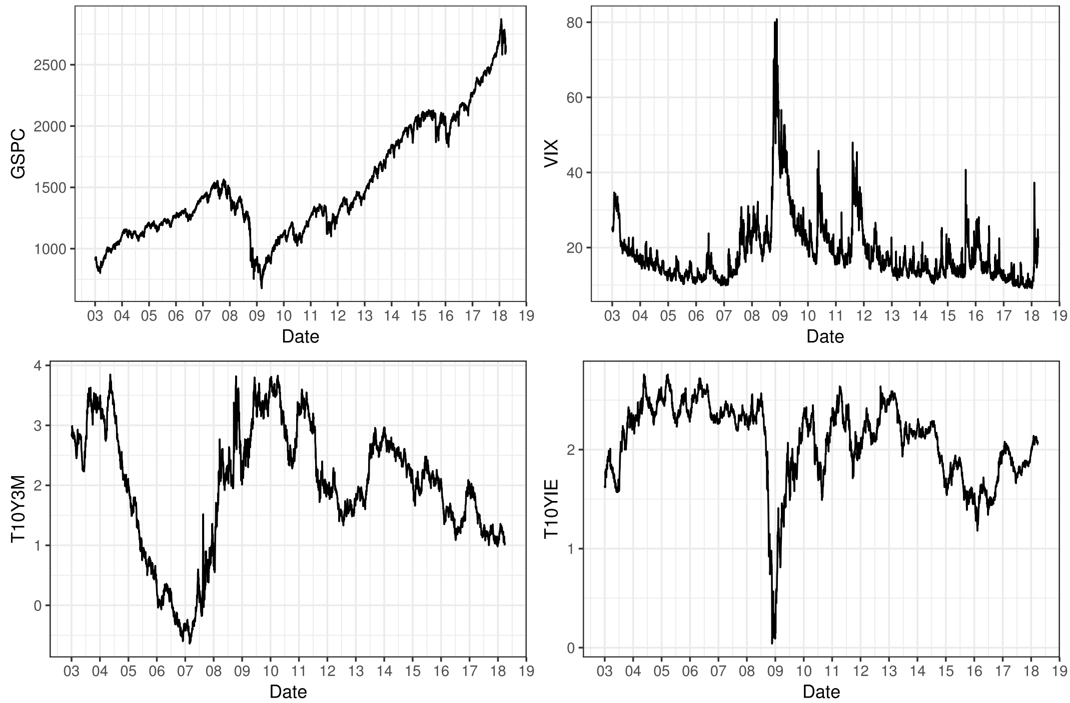

Figure 3.1: Data series from 01-2003 until 04-2018.

Top left: S&P 500, top right: volatility index VIX, bottom left: 10-Year Treasury Constant Maturity Minus 3-Month Treasury Constant Maturity, bottom right: 10-Year Breakeven Inflation Rate.

Before the analysis with the data is started, it is necessary to inspect the data and undergo a process of data tidying. All the data were downloaded from Yahoo Finance and Federal Reserve Economic Data (FRED) database. To fetch the data, R software, together with a well known function `getSymbols` from `quantmod` package was used. The data were downloaded as a time series object which contained a number of observations with non-available (NA) values. All rows which contained purely NAs were deleted. A brief look at the dataset revealed that there are not two consecutive NA rows. After the process was done there were still some observations which contained NAs for some variables. The dataset was refilled using linear interpolation. The figure (3.1) shows the plots of each series.

At the first look, the time series used here are not stationary. The Augmented Dickey-Fuller test proves this statement. Taking the whole dataset from January 2003 until April 2018 the only stationary series based on the ADF test is the volatility index VIX ($p - value = 0.02$). The other series are not stationary, the Breakeven inflation has $p - value$ equal to 0.06, the term spread 0.7 and S&P 500 0.96. However, applying the ADF test for a subset sample from January 2003 until the end of the year 2012 all the time series are assessed as non-stationary. The trading strategy will be based on trading the market thus the GSPC series will be stationarized using logarithmic returns. The data of the 10-Year Breakeven Inflation, term spread or the VIX will be used in levels since economically the level may have a more straightforward explanation and beside that the term spread goes below zero which would make e.g. the return computation cumbersome.

To observe how much the data are stationary one can look at the auto-correlation and partial auto-correlation functions displayed in figure (A.1) in the Appendix A. As we can see all the data exhibit some sort of long memory. The GSPC needs to be stationarized and the dependence occurs only in the first lag. The Break-even inflation and the term spread show dependence in 3 and 4 lags respectively. The VIX indicates a dependency even in the eleventh lag. Looking at the ACF and PACF of the return series one immediately recognize that there are remaining AR or MA structures in the data.

Chapter 4

Methodology

The section about the methodology will be divided into three general parts whereas the first two represent the inputs into the trading strategy which is described in the third section. These inputs are the results of the process of estimating the conditional distribution of stock market returns using linear quantile regression and a model of the GARCH family. For the estimation of quantiles, market data available on the daily bases will be used. Next, a trading strategy, which could possibly outperform the market in the long run, will be developed. The strategy will be built on quantile regression results and after that, it will be compared with the GARCH model. It will be concluded whether the strategy provides useful results in comparison with other trading strategies - for example, the buy-and-hold strategy. The performance of the strategy will be also measured by the Sharpe ratio.

It should be noted here that the goal of the thesis is not to create a manual of how to invest in stocks or the stock market and provide the complete code. Investing in stocks is a very complex task which cannot be systematically successful when one does not account for a large amount of available public information and, to be honest, sometimes the creativeness of a human being is needed. The goal is to show the reader a possibility to enhance his own strategy and incentivize them for further analysis.

4.1 Conditional Distribution Function Estimation

Instead of restricting ourselves to rigid assumptions about the distribution function a non-parametric estimation of conditional distribution function will be used. A fully non-parametric estimation is not possible because of the curse of dimensionality problem which would make the estimation almost unattainable with a larger number of explanatory variables. There are two not fully interchangeable methods for conditional distribution function estimation - distribution regression and quantile regression.

4.1.1 Distribution regression

Suppose the investor wants to estimate the probability that the future return will fall below some specified threshold. This is usually done by Value at Risk methods, which are widely used for a risk management. However, VaRs are not suitable for conditional probability estimation - at best they account for cross-correlation between individual assets in a portfolio. The other way is to use a binary response variable taking the value of 1 if the return falls below the threshold and 0 otherwise. Basically, Foresi & Peracchi (1995) used this procedure, and the estimation in this thesis was supposed to be based on their methods. The problem may be estimated as a logistic regression of the form:

$$\eta(x) = ln\frac{F(y|x)}{1 - F(y|x)} \tag{4.1}$$

where y is the random variable representing the threshold. $F(y|x) = Pr(Y_{t+1} \leq y|X_t = x)$, $\{Y_t\}_0^n$ is the time series of excess returns, $\{X_t\}_0^n$ is a k-dimensional series of explanatory variables and n is the number of observations.

This procedure covers only one part of the distribution. The trader could estimate, that there is a probability of p that the return will be below 0 and $(1-p)$ probability that it is above conditional on some explanatory variables. To model the whole distribution it is needed to repeat the estimation J-number of times and thus create $J + 1$ intervals: $-\inf < y_1 < y_2 < \cdots < y_J < \inf$. These intervals compose together the whole set of possible excess returns since we know that each y_j is a threshold at which the trader computes the probability of not exceeding it.

However, there arises a problem of monotonicity. This problem is illustrated in figure (4.1). The two left plots in the figure show a scatter-plot of the lagged value of the volatility index VIX (a possible explanatory variable in the model) on the x-axis and the return on the S&P500 index on the y-axis (the plot shows only one explanatory variable because of the interpretation simplicity). It is

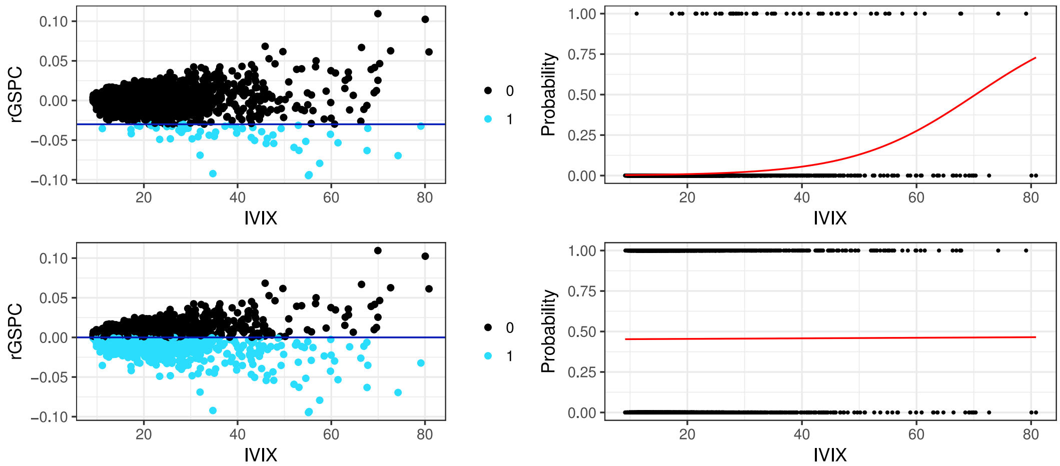

Figure 4.1: Distributional regression monotonicity issue

The two left plots show the dependence of the market returns on the lagged VIX. The plots on the right side provide the results of the logistic regression of the two mentioned variables with different threshold.

immediately obvious that the higher the volatility the more the returns are spread. In both plots, there is a blue line which splits the dataset into two parts. It has the y-intercept equal to -0.03 in the top left plot and equal to 0 in the bottom left plot. Number 1 is assigned to all the cyan points below the line and 0 is assigned to all the black points above the line. The results of the logistic regressions on these data are displayed in the right two plots of the figure (4.1). The top right plot shows that the probability of return falling below -0.03 conditional on the lagged value of the volatility index VIX is close to zero for low volatility but it approaches 0.75 for values of volatility around 80. As you can see, this does not correspond to the top left plot. Moreover, it does not make sense to conclude that when the volatility is high it is almost certain that the return will be less than e.g. -0.03. To illustrate the problem the bottom left plot represents the result of logistic regression when the conditional probability of return falling below zero is estimated. This time the probability that the return will fall below 0 is lower than that it will fall below -0.03 conditional on the high value of the VIX index. One would need to make some adjustments to fulfill the monotonicity but instead of that, it is better to approach the problem with the quantile regression.

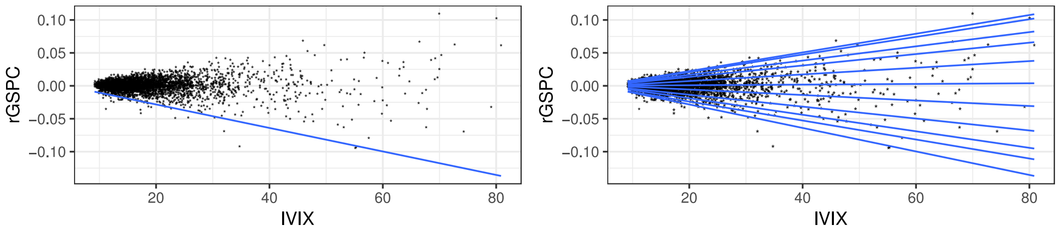

Figure 4.2: Quantile regression example

The points show the dependence of the market returns on the lagged value of the VIX index, blue lines are the results of quantile regressions.

4.1.2 Quantile regression

The quantile regression is very similar to distribution regression but instead of finding the probability of return falling below some specified threshold conditional on explanatory variables one finds the value of return below which the return in the next period falls with previously specified probability.

The general model for quantile regression estimates the following equation:

$$Q_\tau(y_i) = \beta_0(\tau) + \beta_1(\tau)x_{i1} + \cdots + \beta_p(\tau)x_{ip}, \qquad i = 1, ..., n \tag{4.2}$$

The beta coefficients are estimated separately for each quantile. τ's represent individual quantiles and belong to interval $\langle 0; 1\rangle$. The $\beta_j(\tau)'s$ are estimated by minimizing the objective function:

$$Q(\beta_\tau) = \sum_{i:y_i \geq x_i'\beta}^{n} \tau|y_i - x_i'\beta_\tau| + \sum_{i:y_i < x_i'\beta}^{n} (1-\tau)|y_i - x_i'\beta_\tau| \tag{4.3}$$

The essence of quantile regression is in giving different weight to observations above and below the regression line. When one intends to estimate e.g. 1 % quantile the weight given to observations below the regression line is equal to 0.99 and the weight given to observations above the regression line is equal to 0.01.

Nonetheless, it is better to show any model in some example. For simplicity, the same scatterplot of lagged VIX and returns on S&P500 index as in the figure (4.1) is now displayed in the figure (4.2). The result of quantile regression on 1% quantile is represented by the blue solid line in the left plot. The interpretation of this graph states that there is a 1% probability that the return will fall below that line conditional on the lagged value of VIX. It is immediately clear that the relationship for this line is negative - the higher the volatility the more negative is the threshold below which the return will fall with 1% probability. The right plot shows 11 quantile regressions - from the bottom 1%, 2%, 5%, 10%, 25%, 50%, 75%, 90%, 95%, 98%, and 99% quantile. Again for graphical purposes only one explanatory variable was used. Based on these quantile estimates the researcher is able to detect the change in distribution spread. This method is particularly useful in cases similar to the one above because, for example, the classical OLS would not provide us with any information about the true standard error. In situations when the VIX index is low the standard error of market return would be overestimated. On the other hand, in situations with high VIX the standard error would be underestimated.

The quantile regressions have usually different slope coefficients for distinct quantiles. Actually, it is desired to produce different slopes since if all the slopes of quantile regressions were the same, the quantile regression would not be needed. Different slopes mean that the quantile regression lines cross each other. Even though it occurs in every quantile regression it does not harm the model when the cross is for values not present in the data. For example in our case represented in figure (4.2), all the lines cross approximately around zero VIX values - this does not disrupt the model since it can be expected that the volatility index VIX cannot be equal to zero and if so probably different state of the world would surround us and another model would be needed.

The results of quantile regressions are typically plotted as the figure (4.3) shows. Principally the same plots and also other figures related to quantile regression are presented in the article written by Koenker & Hallock (2001) which thoroughly describes an empirical example of quantile regression. In our case, each plot depicts the estimate of the slope coefficient on the y-axis (black dots) corresponding to quantile on the x-axis. The grey area represents 95 confidence bands for quantile regressions. The red solid horizontal line is the OLS estimate surrounded by two dashed red lines which mark the corresponding 95 confidence intervals. The intercept shown in the left plot does not have a useful meaning - it is the estimated conditional quantile function of GSPC return distribution in the situation in which the lagged volatility index VIX (lVIX), term spread (lT10Y3M) and the break-even inflation (lT10YIE) are equal to zero. On the other hand, the next three plots may be interpreted more meaningfully. For example, when the lagged value of volatility index VIX increases by 1 then the return corresponding to 1 % probability of occurrence decreases by approximately 0.002. The quantile regression usually produces

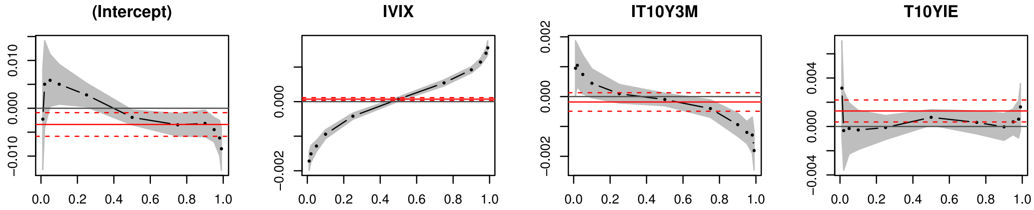

Figure 4.3: Quantile regression summary plots

The plots show the estimates of beta coefficients of explanatory variables from the 11 quantile regressions together with standard errors. The red lines stand for the OLS regression which is surrounded by 95% confidence bands (dashed lines).

these "S-shaped" plots since it captures the dissemination of the conditional distribution function.

4.2 GARCH

The trading strategy will be mainly based on volatility. Thus, a rational consideration would be to compare the final quantile regression based strategy with a well establish volatility model. The best choice is probably the Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) model. The ARCH specification was first introduced by Robert F. Engle in 1982 (Engle 1982) and 4 years later the GARCH model by Tim Bollersev (Bollerslev 1986). During the times it received a lot of attention simply because it can sufficiently predict the value of volatility thus the model copes with heteroscedastic data series and stationarizes the second moment. The model specification looks as follows:

$$\epsilon_t = y_t - x_t'b \tag{4.4}$$

$$\epsilon_t \sim N(0, h_t) \tag{4.5}$$

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} \tag{4.6}$$

where y_t is the dependent variable (usually logarithmic returns of some data time series), x_t a vector of independent variables which are used for predicting the returns or filtering out some information, b a vector of unknown parameters estimated by linear regression - usually the ARIMA model. ϵ_t are residuals from that regression with mean zero and variance/volatility modeled by equation 4.6. The model tries to predict the current value of volatility based on previous p

volatility terms (this is basically a moving average which smooths the volatility predictions) and q innovation terms (usually market returns or the residuals squared) and a constant term. This implies a general GARCH model has $p+q+1$ unknown parameters, however in practice the GARCH(1,1) specification is used very often. It estimates the current value of volatility conditional on previous value of volatility, last squared innovation and a constant term. $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_i \geq 0$.

4.2.1 ARIMA

The GARCH model is usually estimated on residuals from already estimated ARIMA model which consists of three parts. AR is the Auto-regressive part, MA is the Moving-average part and I stands for integrating. If the data are not stationary they need to by first stationarized and then an ARMA model is fitted. The ARMA(p,q) - model is given by:

$$y_t = \mu + \frac{\Theta(L)}{\Phi(L)}\epsilon_t \tag{4.7}$$

$$\begin{gathered} Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \\ \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \ldots - \theta_q \epsilon_{t-q} \end{gathered} \tag{4.8}$$

where Y_t is the data series, L represents the lag operator, ϕ_p are the AR coefficients, θ_q are the MA coefficients, ϵ_t stand for the error term and t is the time index.

4.3 Trading Strategy

4.3.1 Developing

The general goal of any trading strategy is to make large returns with relatively low risk. To achieve this the trader should ideally buy low and sell high. Depending on the trading costs (to avoid excessive trading) they should make this as often as possible. This thesis tries to implement that idea using the quantile regression approach. First of all it is needed to create data which represent the distribution. This is done in the following way: estimate T quantile regressions. For each observation in the dataset create new T observations which represent the fitted values of the dependent variable corresponding to individual quantile regressions. This is done using the procedure in equation (4.2). Because in our case the dependent variable is the return on the market index the fitted values represent return below which the observed return will fall with τ probability. Making sufficiently large number of quantile regressions the fitted values will represent the distribution of stock market returns conditional on a set of independent variables.

Next step is to use these new data for trading. The idea is based on intuition that when the stock market goes down then it is perceived as a buying signal. The opposite is considered to be a selling signal. However, the trader does not have to label each change in price as a signal. The occurrence of a buying (selling) signal could be based on the current return falling below (above) some quantile. For example if the current return is -1 % and the investor wants to buy the market only if the return falls below -2 %, which represents e.g. the 5% quantile then they would do nothing and the signal for the current return would be labeled 0. Without loss of generality the buying signal could be represented by number 1, the selling signal by -1, otherwise 0.

However, it would not be that useful to simply buy the troughs and sell the peaks whenever a signal occurs. The signal may have different strength for different observations. When the market exhibits an uptrend (downtrend), the selling (buying) signal does not need to be strong because it can be expected (based on the momentum) that the market will move higher (lower) in relatively short time period. These trends can be generally represented by a Moving Average (MA) or alternatively by Exponential Moving Average (EMA). When the market exhibits an uptrend the moving average lags behind. Depending on the length of the Exponential Moving Average the longer and the more steeper the market increases the more distant is the current value of the market from the EMA. The whole trading strategy could be based on the connection of these two ideas together. When the market has an uptrend (downtrend) and suddenly moves down (up) it can be considered as a strong buying (selling) signal. On the other hand when the market has an downtrend (uptrend) and drops (increases) even more the investor should not buy (sell) the market. Because of that it is necessary to create a weighting function which assigns the individual signals particular weights.

The weights could be represented by any function which maps the input into a set $\langle 0; 1\rangle$. A suitable candidate is the logistic function. The process is done in the following way: Take a dataset of price movements. Create Exponential Moving Average of length l. Compute the percentage distance of EMA from

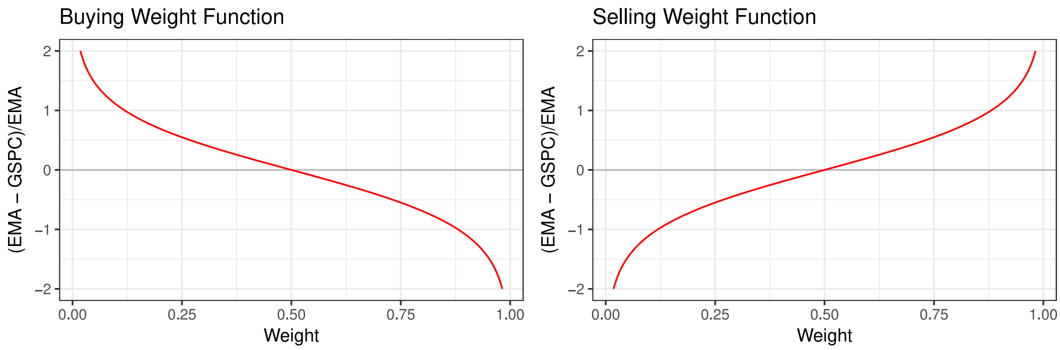

Figure 4.4: Illustration of the relation between the distance of the market price from the EMA and the weight function

the corresponding price. Transform the data using logistic function. Use a logistic function so that the transformation approaches 1 when the difference is significantly negative and zero when it is significantly positive. This procedure assigns each observation in the dataset a weight by which the signal will be multiplied. Let us illustrate an example. When the EMA lags behind the price then the percentage change is negative since EMA minus the price of the market is below zero. The negative change is transformed by the logistic function to a number which is close the 1. By this number the signal is multiplied. If it is a buying (selling) signal which means that the market suddenly dropped (increased) in an uptrend then the signal is multiplied by the computed weight BW_t ($1 - BW_t = SW_t$). The figure 4.4 documents the relationship between the distance of the market price from the exponential moving average and the weight function (the numbers on the y axis as well as the shape of the functions are just illustrative). From the figure it is clear that the more is the market above the EMA the more weight is given to the buying signals and the less weight is given to the selling signals.

The question arises: how much the investor should invest? This thesis makes an assumption that the trader cannot short selling thus going negative in the market and that they cannot borrow additional resources. Then, let us consider they have C_t amount of money in their pocket and they can invest any part of it. The variable C_t is time dependent and for each time it represents the current value of available Cash the investor has. In addition to that the investor has S_t units of the market index. A reasonable assumption is that $S_1 = 0$. The investor buys or sells each time a multiplication of a signal (+1 a buying signal, -1 a selling signal) and the weight.

The investor also does not need to behave the same way in every state of the market volatility. They may behave differently in periods in which the market exhibits higher volatility. It has been observed by the author that in periods of high volatility the market produces a long-term downtrend. This idea can be used and based on the volatility the trader may use a distinct strategy in terms of setting the threshold for the buying and selling signal. It is done by dividing the sample into zones with high and low volatility. The number of zones may be selected arbitrarily. Naturally, more zones can capture more situations.

The question then arises: how to set all the parameters which were mentioned in the previous paragraphs? The quantile below which a buying signal occurs, the quantile above which the selling signal occurs, the combination of both of them. The weighing function shape, its horizontal shifts. The number of zones dividing the market based on volatility and the size of those zones. All of these parameters can be optimized by iteration. After that it may be enhanced by logical reasoning.

4.3.2 Performance evaluation

Absolute returns

The performance of the strategy is firstly represented by the total return it can generate. The absolute return performance will, however, not be measured as the difference between returns on the strategy and on the market. Instead, the percentage difference of the final value from the value of the market will be taken into account. This reflects the comparison to the buy-and-hold strategy.

CAPM and Sharpe Ratio

The Capital Asset Pricing Model developed by (Sharpe 1964) is a basic concept of modern portfolio theory based on the mean of returns as a measure of location and standard deviation as a measure of risk. Intuitively, the higher the risk (measured by the variance) the higher the expected return. In general, the relationship works. However, the mean-variance analysis is reliable only in the situation when the returns are normally distributed or the investors have a quadratic utility function. These characteristics are unfortunately not present in the real world data which implies that the results (build portfolios) are not accurate. Nonetheless, the modern portfolio theory is still used in practice.

How the portfolio is built is described by the figure 4.5. The x-axis and y-axis represent the variance and the expected return of an asset respectively.

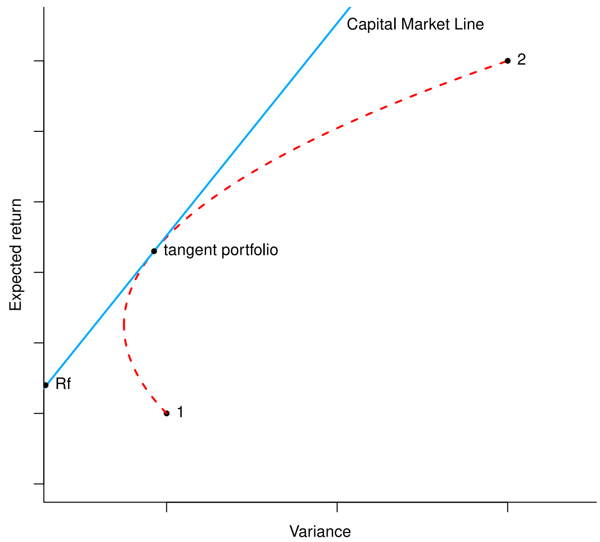

Figure 4.5: Capital Asset Pricing Model

Points labeled as 1 and 2 depict individual assets. An investor is able to get any portfolio on the red dashed line simply by combining asset 1 and 2. The ideal portfolio is represented by the point where the Capital Market Line touches the portfolio curve as a tangent. This point has the highest Sharpe ratio: i.e.

$$Sharpe\ ratio = \frac{r_s - r_f}{\sigma_s} \tag{4.9}$$

where the r_s is the return on the stock (expected return in case of ex-ante analysis and realized return in case of ex-post analysis), r_f is the risk free rate (usually short term treasury bonds) and σ_s is the standard deviation of the stock returns. The Sharpe ratio will be used as performance measure in this thesis.

4.3.3 Trading Costs

Trading costs are usually ignored in studies concerning asset pricing models, portfolio composition or trading strategies. Nonetheless, nothing is free and the same is true for trading. There is no problem to say that a trader could be endowed with a huge amount of assets so that the trading costs can be neglected. The problem, however, appears when one wants to create a systematic trading strategy which outperforms the market. Even a tiny trading cost can completely devastate the result. At this point it can be concluded that it is necessary to include some amount of trading costs - so large that it approximates or even represents real world and so small that it lets open space for active trading.

Next issue arises what costs should one use for historical data and how to predict future costs. It was decided to cope with this problem relatively simply. Asking a question whether the setting of stable trading costs harms the trading strategy results brings us to negative answer. The trading cost are the same for any trading strategy meaning that every strategy will work under the same rules. Most importantly, the active trading strategy is "penalized" for each trade which makes the strategy more realistic. For the purpose of the thesis analysis it was decided to use trading cost set by a company called Interactive Brokers LLC which, according to the authors search, provides the lowest trading costs currently available on the stock market.

Chapter 5

Results

5.1 Quantile regression

At the beginning it is desired to estimate the quantile regressions and check the results. Following regressions are fitted:

$$\begin{aligned} rGSPC_i =& \beta_0(\tau) + \beta_1(\tau) lVIX_i + \beta_2(\tau) lT10Y3M_i + \\ & + \beta_3(\tau) lT10YIE_i + \beta_4(\tau) lrGSPC_i + \epsilon_i \end{aligned} \tag{5.1}$$

where i represents observations, r before the $GSPC$ variable indicates return, l in front of explanatory variables indicates one lag, ϵ is the error term and τ belongs to a set of values: 0.01, 0.02, 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95, 0.98, 0.99. The selection of these values is somewhat arbitrary. The steps in tails of the distribution are smaller to capture the variation. The logic behind these numbers is of that kind that change of 1 % (percentage points) in tails has a much larger effect than the change of 1 % in the center of the distribution. The results of the equation (5.1) are shown in the table (5.1).

The table shows us that the volatility index VIX is significantly negatively correlated with return on GSPC in low quantiles and positively correlated in high quantiles. This suggests that the higher the VIX the more the distribution is spread. For the term spread the relation is similar - the higher the term spread the less the distribution is volatile. The magnitudes for both explanatory variables are very similar - increasing the VIX by 1 decreases the 5% probability quantile return by approximately 0.1 percentage points and the same happens while decreasing the term spread. According to the quantile regression, the Break-even inflation does not have a significant effect in any quantile, thus it will not be used in the development of the trading strategy. The lagged value

	1%	2%	5%	10%	25%	50%
lVIX	-0.002^{***}	-0.002^{***}	-0.001^{***}	-0.001^{***}	-0.0004^{***}	0.0001^{*}
	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.00004)
lT10Y3M	0.001^{**}	0.001^{**}	0.001^{*}	0.0005^{*}	0.0002	-0.0001
	(0.001)	(0.001)	(0.0004)	(0.0003)	(0.0002)	(0.0001)
lT10YIE	0.003	-0.002	-0.001	-0.001	-0.001	0.0003
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.0004)
lrGSPC	-0.095	-0.065	-0.104^{*}	-0.030	-0.023	-0.074^{***}
	(0.071)	(0.089)	(0.055)	(0.045)	(0.030)	(0.020)
Constant	-0.002	0.009	0.009^{***}	0.007^{***}	0.004^{***}	-0.001
	(0.006)	(0.006)	(0.003)	(0.002)	(0.001)	(0.001)
	75%	90%	95%	98%	99%	
lVIX	0.0005^{***}	0.001^{***}	0.001^{***}	0.001^{***}	0.002^{***}	
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	
lT10Y3M	-0.0004^{**}	-0.001^{***}	-0.001^{***}	-0.001^{***}	-0.002^{***}	
	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0004)	
lT10YIE	-0.0004	-0.001	0.0001	0.0001	0.001	
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	
lrGSPC	-0.102^{***}	-0.118^{***}	-0.091^{***}	-0.092^{**}	-0.013	
	(0.028)	(0.034)	(0.034)	(0.043)	(0.052)	
Constant	-0.001	-0.002	-0.003^{**}	-0.005^{**}	-0.008^{***}	
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	
Note:					*p<0.1; **p<0.05; ***p<0.01	

Table 5.1: Quantile regression results

of return on GSPC has in all cases negative slope but significant only in the upper quantiles except for the 99th percentile which is insignificant as well as the lower quantile estimates.

The figure (5.1) shows the scatterplot of pairs of the data. At the first sight, the lagged value of VIX largely influences the distribution of GSPC returns. This is an expected outcome. The figure also shows an interesting result with the term spread and Break-even inflation. The higher is the term spread and the lower the break-even inflation the more volatile are the returns. In case of the term spread it is a bit counter-intuitive since it represents the difference between long-term and short-term interest rates and small difference indicates

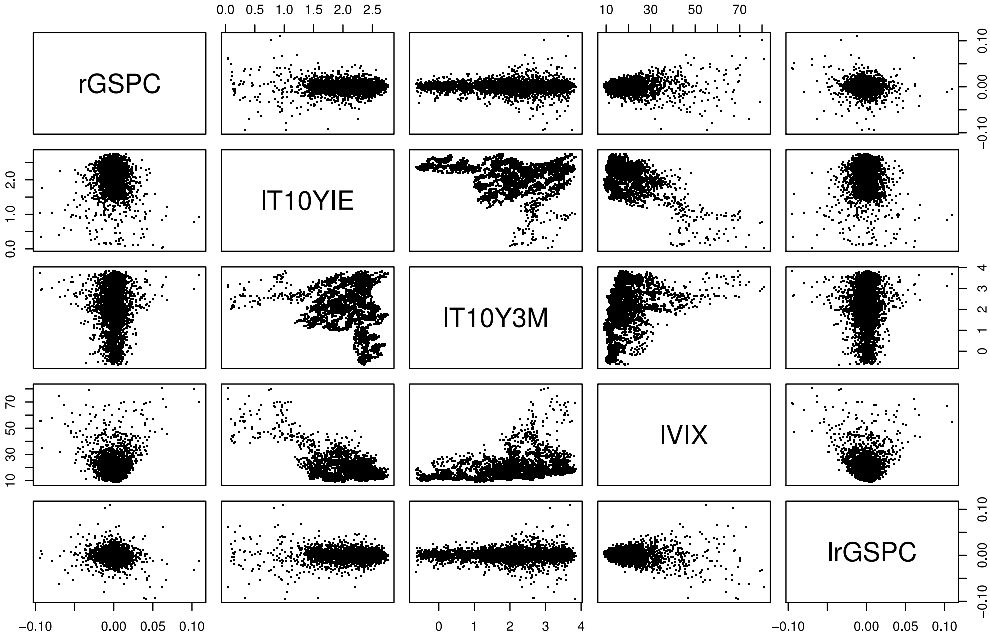

Figure 5.1: Scatterplot of the data

that the market is going through some imbalances. However, a low or even negative difference is just a precursor of the coming recession not a contemporaneous effect of ongoing market slowdown. Even more striking is the plot of returns on GSPC and the lagged Break-even inflation. In the regression which contained all variables and which results are shown in the table (5.1) the break-even inflation did not have any significant effect. Nonetheless, the plot shows a different picture. The lower is the expected inflation the more is the market return spread. This effect could be explained by cross-correlation between the independent variables. The Break-even inflation is strongly correlated with the volatility index VIX. Because of that, the Break-even inflation will not be included in the final regression. The lagged value of the return on GSPC will also not be included since the significant negativeness of the slope coefficients from the quantile regression comes from sampling error - performing the same quantile regression on shorter dataset provides us with insignificant slope coefficients for quantiles which were significant in the regression on the whole dataset.

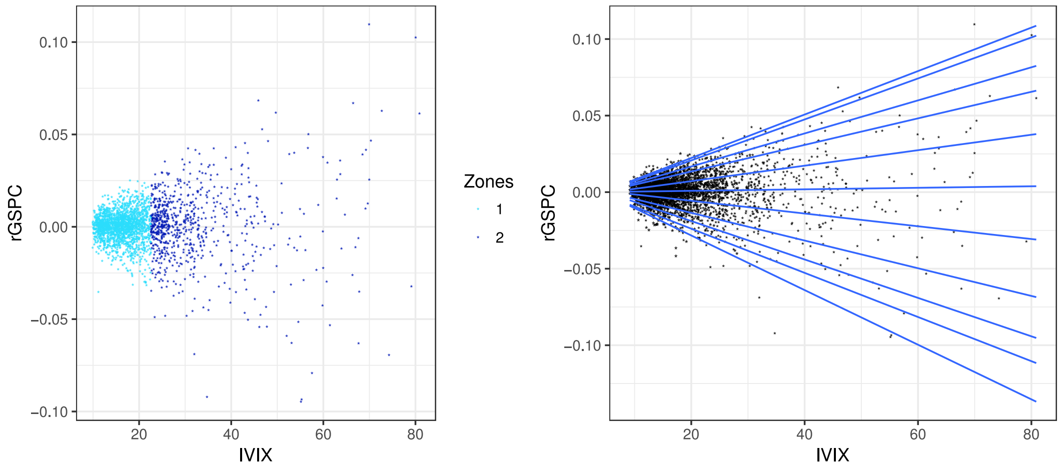

Figure 5.2: Trading zones and quantile regressions

5.2 Trading strategy

First of all, it is necessary to have the right form of data. To achieve this the quantile regressions in equation (5.2) need to be estimated:

$$rGSPC_i = \beta_0(\tau) + \beta_1(\tau)lVIX_i + \beta_2(\tau)lT10Y3M_i + \epsilon_i \tag{5.2}$$

The fitted values are then filled into a data-frame which contains also the date, GSPC prices, logarithmic returns on GSPC and lagged value of the volatility index VIX. The data is then divided into a training and testing sample in the ratio $2:1$. After the creation of the training sample, a new column of zones is created. The dataset will be divided into zones based on the VIX values. The volatility index values range between 9.14 and 80.86. If the data were divided exactly by half (50 % in zone 1 and 50 % in zone 2), it would make the zone with the smaller volatility extremely thin. The median value is 16.23 thus the first zone would include all data points with VIX values ranging from 9.14 to 16.23 and the second zone with VIX values from 16.23 to 80.86 would be approximately 9 times wider. Dividing the sample by the middle value between the maximum and the minimum value of the volatility index VIX would make the second zone sparse. Because of that the division line crosses the VIX range in $2^{-(5/2)} \approx 0.177$. The left plot in figure (5.2) shows the division into zones. The light blue points represent the first zone in which the volatility is smaller and the dark blue points the second zone with higher volatility. This means all observations with VIX between approximately 9.14 and 21.83 belong to the first zone, the rest in zone 2.

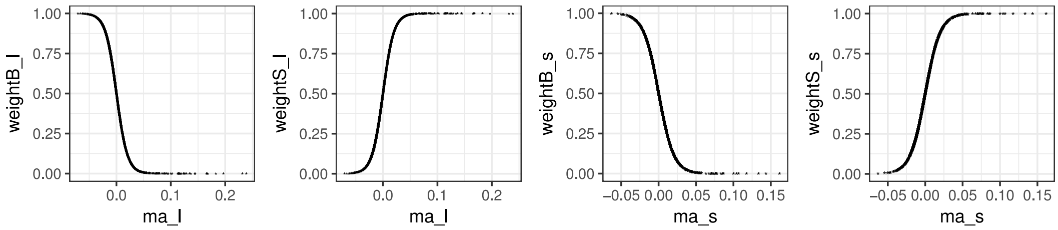

Figure 5.3: Buing and Selling Weights based on Moving Averages

In the next step, moving averages need to be created. Because the trader operates in two zones it would be also suitable to have two different moving averages for the market conditions. The zone 1 represents the market with low volatility thus a shorter exponential moving average is used, particularly 10 day EMA. For the zone 2 with higher volatility a longer 22 day EMA is used. The reasoning behind this idea is that when the market has lower volatility a shorter EMA captures possible changes in the market price. On the other hand when the market exhibits higher volatility the longer EMA "stabilize" its own behavior.

Now two vectors of Exponential Moving Averages were created. The purpose of both of them is just to create the weights. The transformation is done via the following equation:

$$BW_{t,l} = \frac{e^{(-100*EMA_t(l))}}{e^{(-100*EMA_t(l))} + 1} \tag{5.3}$$

where $BW_{t,l}$ represent Buying Weight at time t corresponding to the l which stands for the length of the Exponential Moving Average. The Selling Weight is computed as $SW_{t,l} = 1 - BW_{t,l}$. The EMA is multiplied by 100 to get the percentage changes and to smooth or flatten the logistic transformation.

5.3 Quantile optimization

At this point, we are approaching the iteration part of the thesis. The goal is to find the best combination of quantile thresholds for buying signals and selling signals for zone 1 and 2 which produce the maximum return and/or the maximum Sharpe ratio or some combination of both[1]. Together there

[1]If the trading strategy were doing just a small amount of trades, it may produce very high Sharpe ratio because of the low volatility. This is however not a desired goal.

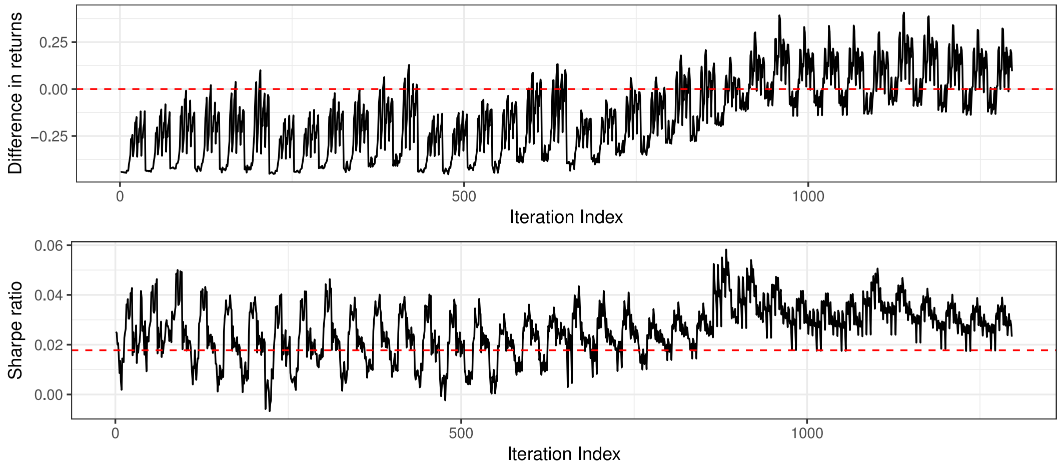

Figure 5.4: Percentage difference in final value of the portfolio and Sharpe ratio

are 11 quantiles, but the buying signal may occur only for the observations below the median and selling signal for observations above the median. The strategy cannot produce both buying and selling signal at once. Thus there are $6^4 = 1296$ possible combinations of buying and selling signals in zone 1 and 2. All of them will be tested. In each step, the buying and selling signals are found. Then they are multiplied by the weights depending on whether it is a buying or selling signal. Cash C_1 is set to 100 and stock S_1 is set to 0. When buying signal at time t occurs and the market operates in zone 1, where the length of EMA is set to 10, a $BW_{t,10}$ portion of cash from period $t-1$ is spent on the purchase of the market and the number of units bought is stored in S_t. This is done row by row for each observation. Of course, the value of S_t changes as the market price goes up or down. At the end, a final difference in returns of the strategy and the market is computed as well as the Sharpe ratio. This procedure is repeated until every combination of buying signal, selling signal and zone is tried.

The figure (5.4) depicts the results of the iteration. The top graph in the figure shows the percentage difference between the final value of the cash and stock the trader has at the end and the value of the market. The red dashed line stands for no difference in final value. The bottom graph represents the Sharpe ratio of each particular iteration. It is immediately obvious that in most cases the strategy does not perform well in terms of absolute returns. On the other hand, the Sharpe ratio is mostly better than for the market which is drawn by the red dashed line - the reason for that is the trader holds a large amount of wealth in cash which does not produce any volatility. In both plots some

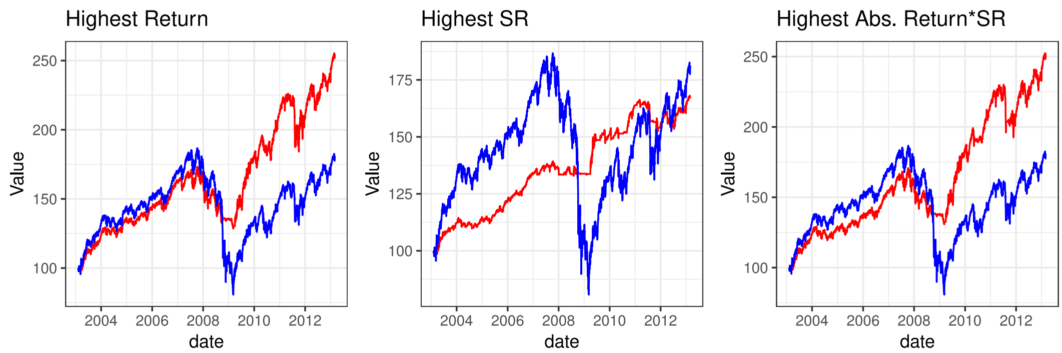

Figure 5.5: Strategies results

Blue line is the market price development. The red line is the value of the cash+stock portfolio.

patterns are present - this is because the iteration changes only one quantile at once. For example the first iteration with index value equal to 1 is the strategy which in both zones 1 and 2 produces a buying signal when the market return falls below the 1 % quantile and selling signal whenever the market return falls above the 50 % quantile of the return distribution based on the VIX and the term spread. The next strategy with iteration value equal to 2 is the strategy which produces in zone 1 and 2 buying signal for market returns below the 1 % quantile but the selling signal in zone 1 is produced for a return above the 50 % quantile and in zone 2 above the 75 % quantile. In the next step, the strategy produces the selling signal in zone 2 above the 90th percentile, in the next step above the 95th etc. When it reaches 99th percentile for selling signals in zone 2 then in the next step the buying signal in zone 2 is produced below 2 % quantile and the selling signal above the 50th percentile. This procedure ends when all combinations are tried - after 6^4 trials. The pattern is created because e.g. in first $216 = 6^3$ iterations the buying signal in zone 1 is produced for return below 1 % quantile and only the other signals change then for other 216 iterations it is produced for returns below 2 % quantile. The selling signal for zone 1 changes only once in 36 iterations and the buying signal in zone 2 changes once in 6 iterations.

Now it is time to look at the best result. The figure (5.5) displays three situations: The strategy with the highest absolute return at the end of the period is depicted in the left plot, the middle plot presents the strategy with the highest Sharpe ratio and the left plot shows the strategy which produced the best multiple of absolute return and Sharpe ratio.

Highest Return

The strategy with the highest return behaves in the following way: if the return falls below the median return[2] and the market exhibits low volatility (zone 1), then it produces a buying signal. If the return falls above the 75th percentile during low volatility, then it produces a selling signal. For zone 2 the thresholds are 10th percentile - buying signal and 98th percentile - selling signal. There are together 965 buying signals and 433 selling signals. It means that there is approximately one trade per 1.8 days. The strategy has 40 percent higher value at the end than the market and it was able to produce it with Sharpe ratio almost 3 times higher than has the market ($SR = 0.0436$).

Highest SR

This strategy is more conservative. It does not produce excessive returns. When the market increases the returns are significantly lower. However, this strategy does not lead to large losses during the crisis. The setup is following: buying signal zone 1 - 25th percentile, selling signal zone 1 - 50th percentile, buying signal zone 2 - 10th percentile, selling signal zone 2 - 50th percentile. It is obvious here the strategy produces sell more often than buy. There are only 530 buying signals whereas 1273 selling signals. Together it translates into one trade per 1.4 days. The Share ratio is equal to 0.058 - approximately 3.28 times higher than the one of the market.

Highest Abs. Return*SR

This strategy is a combination of the previous two and produces 1.39 times higher value than the market does with Sharpe ratio around 0.047. This strategy does not produce so large drop during the crisis as the first strategy does but it also more lags the market during an uptrend. The setup is as follows: buying singal zone 1 - 25th percentile, selling signal zone 2 - 90th percentile, buying signal zone 2 - 10th, selling signal zone 2 - 95th percentile. This strategy buys more times in both zones than it sells but generally it does not trade as much. There are 530 buying signals and 201 selling signals which totals in one trade per 3.47 days.

5.4 Weight Function Shifts

The trading strategies produced in the previous section were developed using iterations only through the individual quantiles and zones. However, the weight

[2] It practically means whenever the return is negative

function could be also changed. In this case, the weight functions for both exponential moving averages will be shifted horizontally and again the best strategy according to the absolute return and Sharpe ratio will be selected. At this point, the first two particular strategies produced in the preceding section will be used. In other words, only the weights for each signal will change. In the section (5.3) the weights were given by equation (5.3). The weight function was build in such a way that when the current market price was equal to the Exponential Moving Average of length depending on the zone then the assigned weight to either buy or sell signal equaled 0.5. In this part, the shifting of the weight function will give more importance to buy or sell signals depending on the direction of the shift.

The setup of the iterations looks as follows:

$$BW_{t,l=10,r} = \frac{exp(-100 * (EMA_t(10) + r))}{exp(-100 * (EMA_t(10) + r)) + 1} \tag{5.4}$$

$$BW_{t,l=22,s} = \frac{exp(-100 * (EMA_t(22) + s))}{exp(-100 * (EMA_t(22) + s)) + 1} \tag{5.5}$$

where t is the length of the dataset, l represents the length of EMAs and $s, r \in a_n; a_1 = -0.04; a_{n+1} = a_n + 0.002; n = 1, 2, \ldots, 41$. The sequence from -0.04 to 0.04 with steps by 0.002 is selected arbitrarily since it can be expected that if the function were shifted too far, then all the weights given to either buy or sell signals would equal to zero and 1 or the other way around. This would mean that the strategy would only buy at the beginning or not even start buying. It is needed to say that every combination of r and s is examined and it is done separately for the strategy with the best absolute return and with the best Sharpe ratio. The figure (5.6) displays the result. The bigger red points in each plot represent the original strategy without any weight function shifts.

The two graphs on the left side show the performance of the tradings strategy which had the highest return in the previous section while shifting the weight function. Because the red point in the first graph is near the top of all points it is obvious that shifting the weight function did not add any further return. Particularly, shifting the weight function of the slow moving average

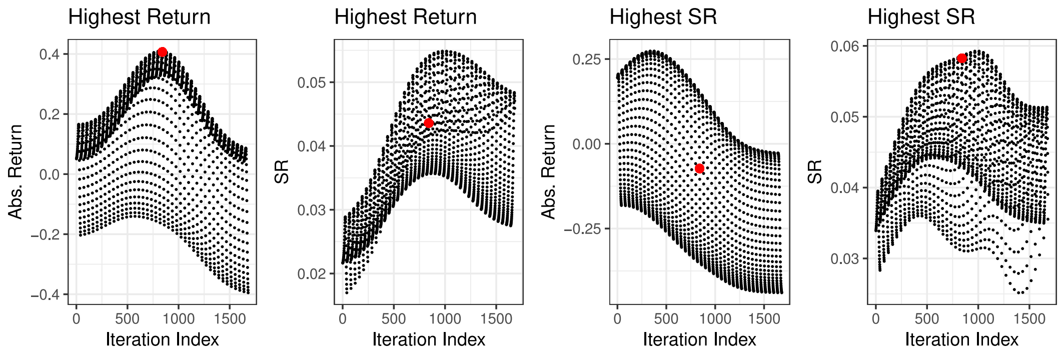

Figure 5.6: Weight function shifts

by 0.002 to the right adds a small amount of return but decreases the Sharpe ratio.

On the other hand, the right plot shows that shifting the weight function increases the Sharpe ratio of already the best Sharpe ratio strategy from the previous section iteration. Nonetheless, it improves it from 0.058 to 0.059 which is not much and it rapidly decreases the final absolute return from - 7 % to - 23 %.

Already from these conclusions, it is conspicuous that there may be some negative Sharpe ratio / Absolute return relationship. To clarify it a scatter plot is produced and shown in figure (5.7). The red dots represent strategies which are produced from the Best Sharpe ratio setup and black dots stand for strategies produced from the Best Absolute return setup. While all the black points are shifted to the right region where are the higher returns, the red points are more extended to the upper part of the plot with higher Sharpe ratio values. The "efficient line" is naturally composed of points more close to the top right corner. It is important to note that until now only two strategies were analyzed in detail. There can be another strategy which did not produce the best Absolute return or the highest Sharpe ratio but iterating over weights it may produce even better combination of both - the Sharpe ratio and Absolute return. The thesis will cope with this topic in the discussion chapter. We may conclude that the iteration over weights does not bring particularly better results when the same performance indicator is taken into account.

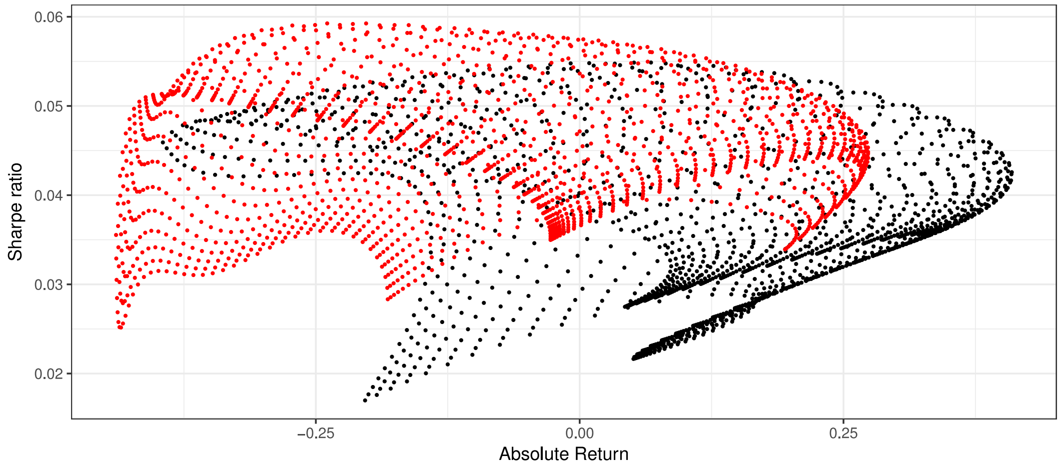

Figure 5.7: Weight function shift results

The relationship between the Sharpe ratio and the Absolute return with different weight function settings. Red dots represent the originally best Sharpe ratio strategy and the black dots the originally best Absolute return strategy.

5.5 GARCH comparison

The section (5.3), Quantile optimization, estimated the distribution of market returns based on the volatility index VIX and the term spread. By a sufficient number of quantile curves, it produced the estimates of the conditional distribution of market returns. In this section the GARCH model will estimate the shape of the distribution, of course, assuming the normality of return.

	mu	ar1	ar2	ar3	ma1	ma2	ma3	omega	alpha1	beta1
ARMA(3,3) GARCH(1,1)	0.0006	-0.30	-0.32	0.57	0.24	0.25	-0.65	0.0000	0.0828	0.9042
GARCH(1,1)	0.0005							0.0000	0.0839	0.9029

Table 5.2: GARCH model coefficients

First of all, it is needed to set a proper model for GARCH which means fitting ARMA model to the logarithmic returns and then the GARCH specification on the residuals. 36 ARMA(p,q); $p, q \in \{1, 2, 3, 4, 5, 6\}$ models were tested and the one with the lowest Akaike Information Criterion (AIC) was selected. The lowest AIC had ARMA(3,3). Overall, two GARCH models were examined - one with ARMA residuals and one directly on the returns. The results of both models are shown in table (5.2). The coefficients from the GARCH part of both models differ in the third or fourth decimal place. This brings us to the simplification that GARCH(1,1) will be directly estimated on logarithmic returns since filtering the auto-regressive or moving-average part does not bring significant additional value.

The procedure looks as follows: Eleven columns of inverse cumulative distribution function corresponding to preselected quantiles were created. Each row of new observations was multiplied by the particular value of volatility estimate. The equation (5.6) describes this procedure.

$$\hat{q}_t(\tau) = \Phi^{-1}(\tau) * \hat{h}_t + \tilde{r} \tag{5.6}$$

Φ is the normal cumulative distribution function, t is the length of the dataset, $\tau \in \{1, 2, 5, 10, 25, 50, 75, 90, 95, 98, 99\}$, $\hat{h}_t$ are the fitted values from GARCH(1,1) model and $\tilde{r}$ is the median return on the market. Then the same zones as in the quantile regression approach were used as well as the exponential moving averages and the weight functions. The figure (5.8) represents the results of GARCH variance estimates together with realized market returns and market returns divided by the estimated volatility. In the middle plot of the figure, it is seen that the market returns have stable mean but that the second moments exhibit the volatility clustering phenomenon. Dividing in by the GARCH fitted values of volatility it produces a series which on the first look has a stable variance. This is confirmed by the Auto-correlation plots in figure (5.9). The left plot represents the ACF for squared market returns and the right plot the ACF for squared market returns divided by the GARCH estimated variance.

The next figure (5.10) shows the quantile returns corresponding to GARCH fitted values plotted with the lagged value of the VIX index. It is clearly visible that the higher the yesterday's value of VIX the more is the distribution of returns spread. The black line represents the 1% quantile, red 10% quantile and blue 25 % quantile. Generally, the relationship between GARCH volatility estimates and VIX index is linear as shows the figure (5.11).

The results of the same strategies but with GARCH volatility estimates are shown in figure (5.12). The process of iteration is not that satisfactory as it was with quantile estimates. In a smaller amount of cases, the final percentage difference in value was higher than 0 and none of the models exceeded 30 % value difference. On the other hand, the Sharpe ratio is most of the time above the market but the reason for that is poorer return performance.

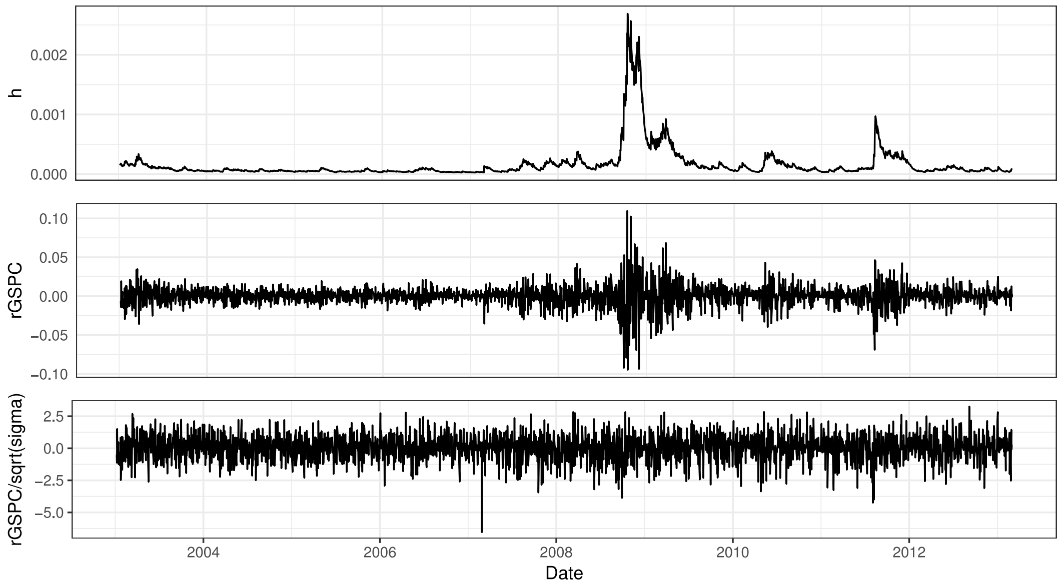

Figure 5.8: GARCH fit results

Top plot: estimated variance $\hat{h}_t^2$, Middle plot: market logarithmic returns, Bottom plot: Division of market returns by the square root of variance estimate

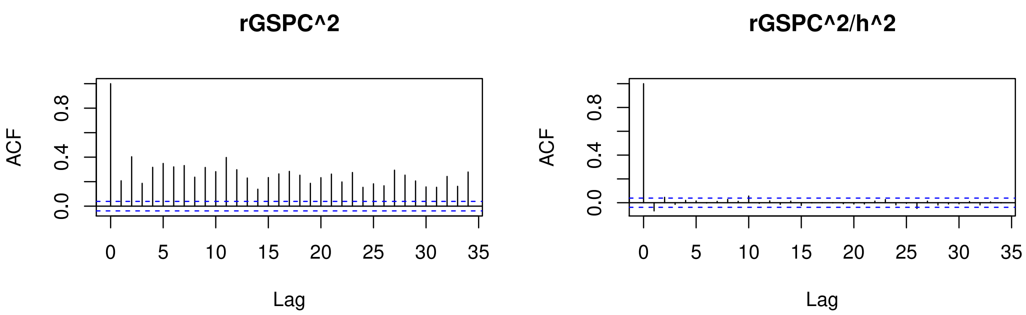

Figure 5.9: ACF of squared returns and ACF of squared return divided by GARCH estimated variance

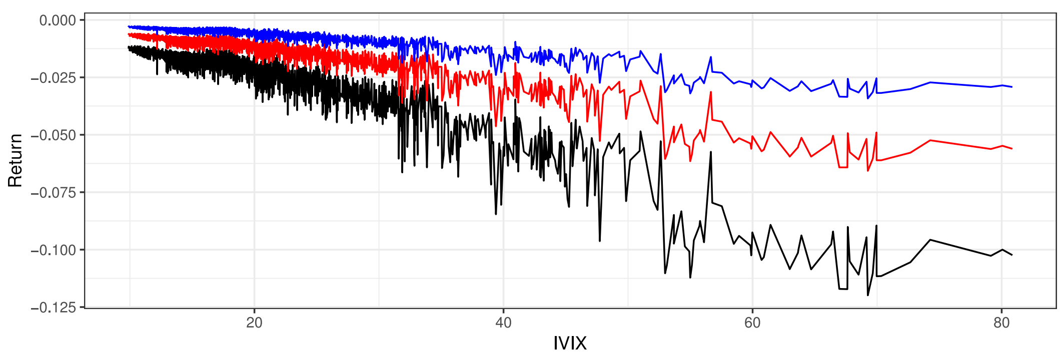

Figure 5.10: GARCH fitted quantile returns with the lagged VIX

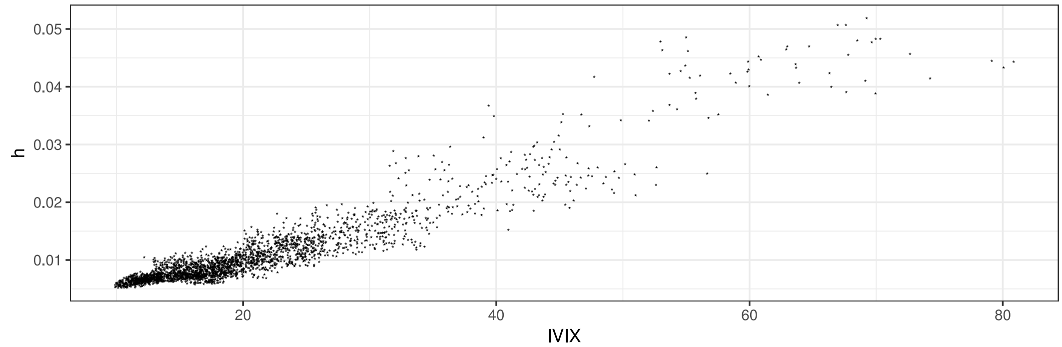

Figure 5.11: *GARCH fitted volatility estimates with the lagged VIX*

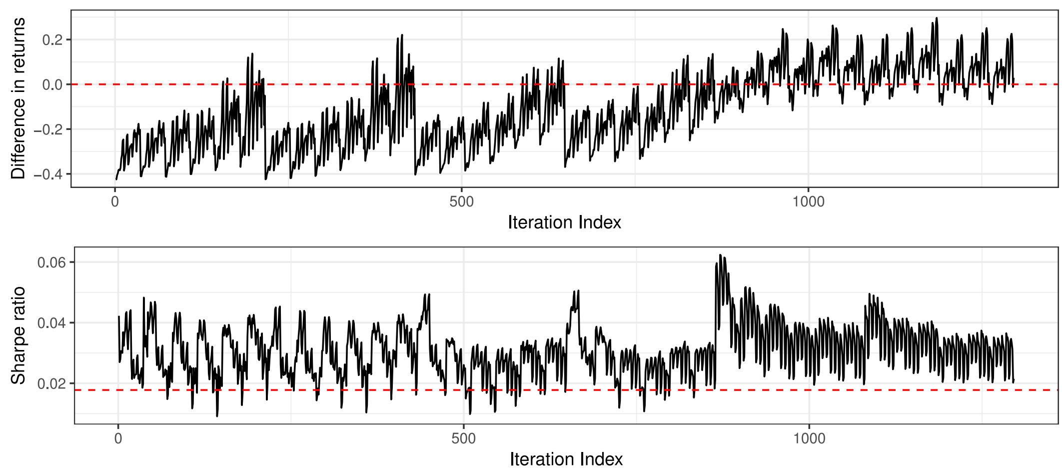

Figure 5.12: *Percentage difference in final value of the portfolio and Sharpe ratio based on GARCH volatility*

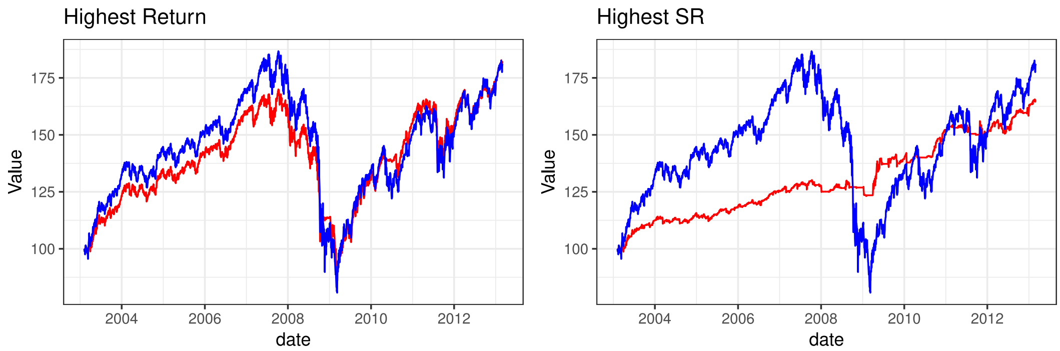

Figure 5.13: GARCH based distribution, quantile regression strategies

The two best strategies from section 5.3 (one with the highest absolute return and one with the highest Sharpe ratio) are now tested but with GARCH distribution estimates. The figure (5.13) displays the development of the strategy which had the highest absolute return using quantile regression in the left plot and the development of the strategy which had the highest Sharpe ratio in quantile regression in the right plot. It can be seen that in this particular example the GARCH model does not add much value. The return based strategy basically imitates the market movements and most of the time the value of the portfolio consisted of Cash C_t and the market units S_t is below the market value. On the other hand, the Sharpe ratio is relatively higher - 0.025 compared to 0.017 which has the market. The second strategy which had the highest Sharpe ratio in quantile regression framework performs rather well. The final return is below zero but the portfolio value does not drop significantly over recession/crisis periods. The Sharpe ratio is substantially higher comparing it with the market. The value 0.056 is 3.17 times higher than the market Sharpe ratio (0.017).

Best GARCH based strategy

The best performing GARCH based strategies are depicted in figure (5.14). The left plot shows the development of the value of portfolio which is based on the strategy maximizing the absolute return. The setup does not account for the zones - it buys the market whenever the return falls below the median return and sells above 90th quantile. Out of 2560 days is produced a signal in 1467 of them.

The GARCH based strategy with the highest Sharpe ratio is similar to the quantile-based. It also produces sell signals whenever the return falls above the

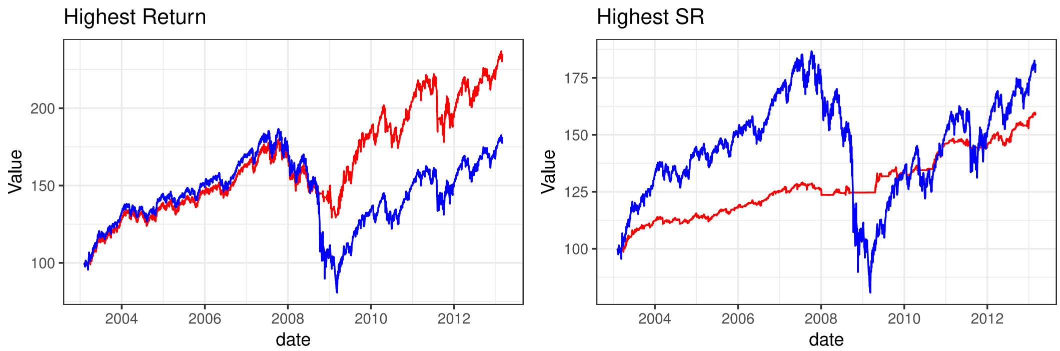

Figure 5.14: Best GARCH based strategies

median. The buy signal for zone 1 is also the same (25th quantile), however, in zone 2, it buys in fewer cases - particularly below 2nd quantile.

All the strategies, their settings and performances are shown in the table (5.3). The majority of them significantly differ in term of absolute return as well as the Sharpe ratio. The quantile regression based strategies produced higher returns regardless of the metrics which was maximized.

				Quantile signal			
				Zone 1		Zone 2	
	Absolute Return	1 trade in days	SR	B	S	B	S
QR Return	40 %	1.8	0.044	50	75	10	98
QR SR	-7 %	1.4	0.058	25	50	10	50
QR MIX	39 %	3.5	0.047	25	90	10	95
$GARCH_{QR}$ Return	0 %	1.9	0.025	50	75	10	98
$GARCH_{QR}$ SR	-8 %	1.4	0.056	25	50	10	50
GARCH Return	29 %	1.7	0.036	50	90	50	90
GARCH SR	-12 %	1.5	0.062	25	50	2	50

Table 5.3: Strategies performance summary

5.6 Out-of-Sample

Until this point, every part of the analysis was done in-sample. All the models were estimated, examined and tested on one dataset. Now, each of the previous models will be evaluated on the testing sample. All the data used in previous analysis started in January[3] 2003 and ended in March 2013. For the out-of-sample analysis, the remaining part of the dataset will be used - the period from March 2013 until April 2018.

[3] After the Exponential Moving Averages were computed the NA rows were deleted which reduced the sample by one month.

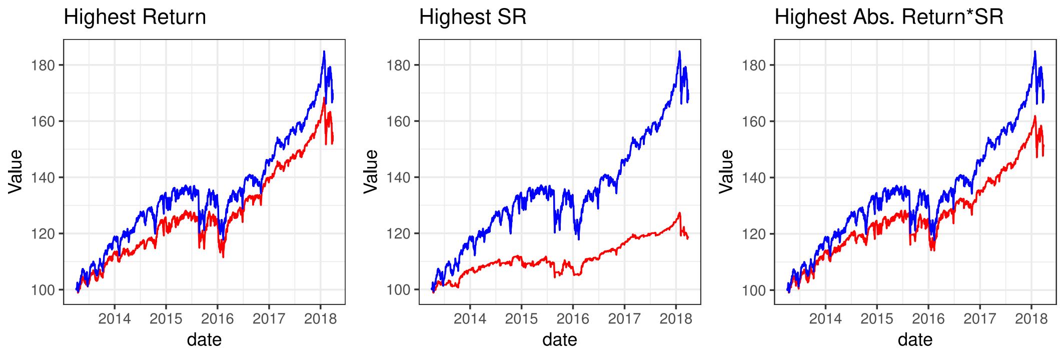

Figure 5.15: Out-of-Sample QR based strategies

The out-of-sample performance of the quantile-based strategy is depicted in the figure (5.15). The models are the same as they were in the in-sample analysis. On the first look, none of the strategies performs well in out-of-sample. The absolute returns are below the market in any case. In this case, the market Sharpe ratio is equal to 0.052. The final value of the portfolio created by the strategy which maximized the return in-sample is 8 % below the market and the Sharpe ratio is just about one thousandth higher. There were 939 trades in the period meaning approximately one trade per 1.34 day. The second strategy which originally maximized the Sharpe ratio now performs worse than the market even in term of the Sharpe ratio which is equal to 0.041 (77 % of the market SR). The third strategy composed of a combination of the previous two produced also unsatisfying results. The final portfolio value is about 11 % lower compared to the market value and SR is equal to 0.053.

Speaking about GARCH Out-of-Sample the results are analogous and plotted in figure (5.16). The model which maximized the GARCH based strategy returns (left plot in the figure) produces lower final value as well as lower Sharpe ratio (value difference: - 10 %, SR: 0.048). Totally, it produces one trade per 1.75 days. The second strategy (right plot in figure 5.16) which optimized the Sharpe ratio produces both lower final portfolio value as well as lower Sharpe ratio (value difference: - 35 %, SR: 0.029).

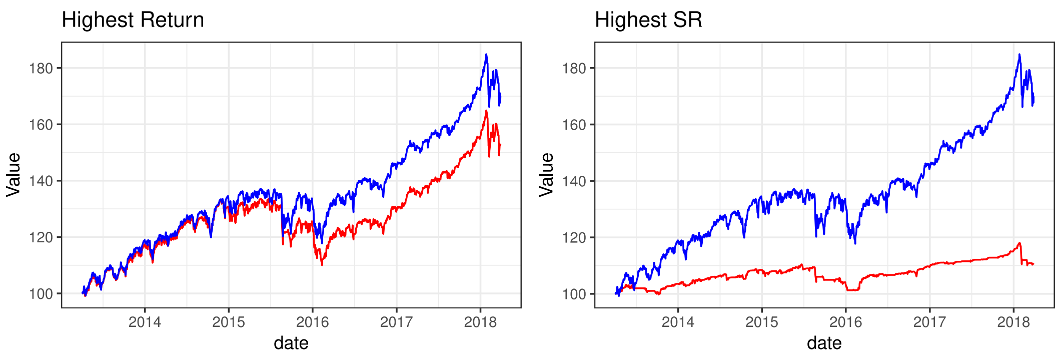

Figure 5.16: Out-of-Sample GARCH based strategies

Chapter 6

Discussion

The results of the trading strategies in the previous section are not that satisfactory as it may be desired, especially considering the out-of-sample performance. A couple of important remarks need to be mentioned about the strategy development process. From the beginning until the end of the analysis a great number of values were selected arbitrarily. This means that the thesis left many unfilled gaps and potentially unanswered questions.

Starting with the explanatory variables - instead of the volatility index, term spread or Break-even inflation, different market variables could be taken into account. The general advantage of quantile regression is that it can capture the behavior of the dependent variable distribution conditional on a set of explanatory variables which is not the case of, for example, the GARCH model because it takes into account only the market returns. That is also the reason the quantile regression was used in this thesis - the spread of the distribution can be examined with the help of other explanatory variables. Secondly, another set of quantiles could have been chosen. 11 quantiles can relatively precisely estimate the distribution shape but in this case, it is still unknown, how the distribution looks e.g. between 25th and 50th quantile and whether estimating another quantile regressions may bring some positive results.

The most arbitrary parts of the thesis are the number of zones in which the sample was divided, the threshold between the zones and the variable on which the division was based. The reason why 2 zones were selected was because of more intuitive interpretation between low and high volatility regions. It would be probably better to use for example 3 to 5 zones which would capture more precisely the different market conditions. The zones were also selected based on only 1 dependent variable - the volatility index VIX. However, the quantile regressions were estimated not only with the VIX index but also with the term spread which could also serve as a "zone distributor" - adding more zones also based on this variable may be potentially beneficial. Nevertheless, dividing the sample based on another variable would double the number of zones. Four zones in our case would mean 6^8 possible outcomes which are almost 1.7 million. This is a very large number of possibilities. Of course there is no need to test all of them but still, this amount exponentially grows with every new element. The reason why the zones were based on volatility index VIX is mainly from the motive that it is a forward-looking indicator which is computed from the expectation of the S&P500 index and it does not require a large amount of iterations.

Next, the weight function may have been also optimized in a better way. It does not need to be only shifted but it may be "stretched" to both sides or even have a completely different specification. The weight function served the goal of telling the investor whether the current market drop is a good chance to buy the market or whether the market has changed the trend and they should wait. The weight function was also based on the exponential moving averages which may have been set to completely different length. However, if we wanted to test all the combinations of moving averages, weight functions, quantile thresholds, zones and variables in zones the solution would be practically unattainable via iterations.

This brings us to the question of how to achieve better optimization but relaxing the iteration part. If we wanted to stick with the quantile regression and all its features it would be probably needed to define a cost function which would measure the loss or distance from the optimum or more precisely it would tell us how should be the variables changed in order to achieve better performance. However, this setup would need a training dataset from which it would learn how to proceed and how to set all the parameters. With this step, we are basically approaching the machine learning algorithms - more particularly the neural networks which are based on a similar idea. Nonetheless, the neural networks are black boxes which do not uncover the model specification. Besides that, the training dataset had to be imported - this means to have some already working trading strategy or some other type of data from which the investor could extract the information of how much and when to buy or sell.

Coming back to the developed strategy a great amount of information and its pattern can be extracted from the figure (5.4). In the first half of the iterations, the strategies do not perform very well in term of absolute return and the Sharpe ratio is more volatile. In this part of the plot, the buying signals for zone 1 are set too low and the strategy basically tells the investor to hold onto a large portion of cash. In case of the lower frequency pattern (selling signals in zone 2) - it does not produce much change to the final value since the pattern is quite flat. This tells us that the selling signal is not particularly decisive in zone 1.

To see what drives the performance of the trading strategy, let us look at the figure (6.1) which shows 25 best performing strategies (from the left) based on absolute return and quantiles for which a signal was produced. The crosses represent the zone 1 with lower volatility and the circles the zone 2 with higher volatility. Even though none of the strategies produced selling and buying signal at once for the median the signals are colored - buying signals in blue and cyan and selling signals in black and red. Taking into account the results the best strategies bought relatively often in low volatility producing the signal for 25th and 50th quantile. This is a reasonable conclusion since when the market exhibits low volatility the prices usually increase and the trader wants to be in the market. Beside that, it produces selling signals also relatively often even though not for the median but for the 75th and 90th quantile. This helps the trader to make cash available for next purchases. Analyzing the second zone with higher volatility the signals were produced more rarely. Surprisingly, the first 20 strategies produced buying signal for the 10th quantile only. The selling signals are produced for the 95th and 98th quantile. This generally means that when the market exhibits higher volatility then the trader should buy and sell quite seldom and only if there is a strong signal represented by the actual return falling in the tails of the return distribution. These strategies outperformed the market by more than 28 % during the analysis period.

The figure (6.2) shows the 25 best performing strategies in terms of the Sharpe ratio. The buying signals are produced similarly as in the best absolute return strategies - in the zone with low volatility it is below the 25th quantile and for the second zone, it is below 10th or 5th quantile in most of the first 25 strategies. Generally, it is just a slightly smaller amount of buying signals for the best Sharpe ratio strategies. On the other hand, the selling signals were produced much more often - above the 50th and 75th quantile. This indicates that the trader should sell practically whenever the market exhibits a return above the median. Another point is that the best Sharpe ratio strategies are

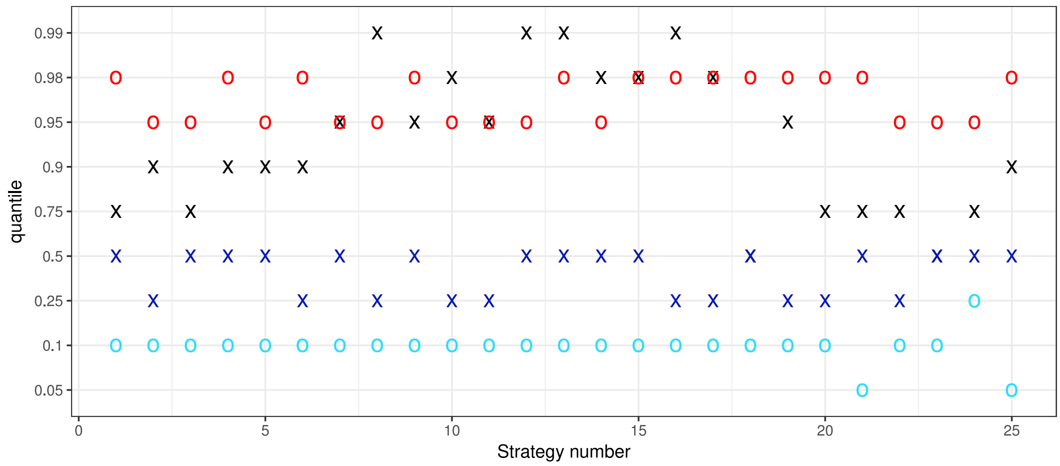

Figure 6.1: 25 best performing strategies (Absolute return)

Quantiles below or above which the buying and selling signals were produced. Crosses represent the zone 1 and circles the zone 2. Buying signals are in blue and cyan color and selling signals are in black and red color.

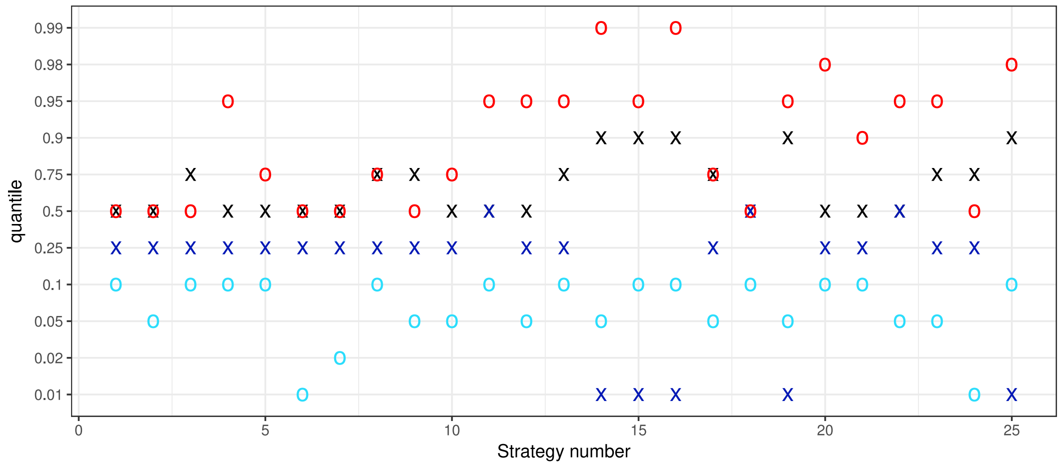

Figure 6.2: 25 best performing strategies (Sharpe ratio)

not so consistent as the return maximizing strategies since the rules when to buy or sell vary through the whole distribution even in the first 25 strategies.

Even though the signals are the nodes which decide when and what action to take, the weight functions and exponential moving averages determined the amount bought or sold. It holds that the longer the moving average the less the current price movement crosses the EMA line and the more time it takes the model to switch the regime. Longer moving averages would skip smaller "waves" in prices and the trader would miss the opportunity to buy or sell. Still, it is needed to say that the prices move above or below the moving average relatively quickly - if the price had a stable pace of growth for ten days and then it had the same pace but decreasing the 10-day moving average would be above the market price just in three days. Accounting for the empirical observations that market slowdowns are usually faster and last shorter the market price would be below the moving average even faster. Generally, longer moving averages catch more the global market conditions and may be used for recognition of the market recession or crisis whereas the shorter ones detect short-term market movement which can be used for active trading - this is the case of 10 and 22-day moving averages.

The results show that the GARCH model does not provide better results in comparison with the quantile regression. This, however, may be only a case-specific conclusion. Particularly, different weight function could make the GARCH model better. Both methods, the quantile regression, and GARCH model try to recognize the same characteristic but from a distinct perspective. The GARCH takes into account only the historical returns and based on that it estimates the volatility. Quantile regression estimates also the volatility but it can do so conditional on a broad set of market data which makes it a bit superior even because it may include squared returns as well as the GARCH.

The reason why the out-of-sample performance was not satisfactory may be because the model was developed on data which included the stock market crisis in 2008. Based on that the strategy could be potentially beneficial in protecting the investor from large losses which did not occur in the testing period.

Taking into account the trading cost it would not change dramatically the results. In all cases, the strategy produced signals from 700 to 2000 times during the 2560 days long period. This translates to approximately 70 to 200 trades per year. Based on the trading costs of Interactive Brokers LLC one trade costs 1 US dollar or \$0.005 per share (whichever is higher). Buying a stock which shares cost \$100 then if the investor wanted to buy shares in the value of \$20000 it would cost him only \$1 or in other words 0.005 %. 200 trades per year make total costs equal to 1 %.

Chapter 7

Conclusion

This thesis represents an attempt to develop a trading strategy which could outperform the stock market in a very long run. It uses the quantile regression framework and GARCH models to predict the conditional distribution function of stock market returns. In case of quantile regression, it uses the volatility index VIX and the term spread represented by the difference between 10-year minus 3-month treasury constant maturities as the explanatory variables. GARCH(1,1) model is used for distribution estimates in the second case.

Based on the estimated distributions buying and selling signals for each day are produced when the daily return falls below or above predefined quantile of the return distribution. The buying signals are produced in the lower half of the distribution and the selling signals in the upper part. The signals are then multiplied by a weight function which is based on 10 and 22-day exponential moving averages and gives more weight to buying (selling) signals which are produced in an uptrend (downtrend) and less weight to buying (selling) signals in a downtrend (uptrend). This helps the model to stop buying in a market recession and not sell when the market increases.

The whole model was developed and tested on the S&P 500 logarithmic returns from January 2003 until the end of the year 2012. The out-of-sample performance was tested on the same series from the end of 2012 until April 2018.

The thesis developed a trading strategy which outperformed the market in-sample in terms of absolute return as well as in terms of the Sharpe ratio. The strategy performed well because it was able to detect the recession/crisis environment which reduced large losses. However, it did not perform better out-of-sample - this was probably caused by the selection of the testing sample which did not include a proper crisis. The thesis brought some interesting ideas which could be further extended since a large number of parameters in the model were selected arbitrarily using only logical reasoning. Finding a better way for optimization the performance of the trading strategy may be enhanced and the trader could outperform the market even in an uptrend.

Bibliography

Amisano, G. & R. Giacomini (2007): "Comparing density forecasts via weighted likelihood ratio tests." *Journal of Business & Economic Statistics* **25(2)**: pp. 177–190.

Azoff, E. M. (1994): *Neural Network Time Series Forecasting of Financial Markets*. New York, NY, USA: John Wiley & Sons, Inc., 1st edition.

Bernstein, P. L. (1987): "Liquidity, stock markets, and market makers." *Financial Management* **16(2)**: pp. 54–62.

Bollerslev, T. (1986): "Generalized autoregressive conditional heteroskedasticity." *Journal of Econometrics* **31(3)**: pp. 307 – 327.

Bollerslev, T., R. Y. Chou, & K. F. Kroner (1992): "Arch modeling in finance: A review of the theory and empirical evidence." *Journal of econometrics* **52(1-2)**: pp. 5–59.

CBOE VIX (2018): Available at `http://www.cboe.com/vix`.

Cenesizoglu, T. & A. Timmermann (2007): "Predictability of stock returns: a quantile regression approach." *Technical report*, Technical Report, Cirano.

Cenesizoglu, T. & A. Timmermann (2008): "Is the distribution of stock returns predictable?" .

Chordia, T. & L. Shivakumar (2002): "Momentum, business cycle, and time-varying expected returns." *The Journal of Finance* **57(2)**: pp. 985–1019.

Cochrane, J. H. (2008): "State-space vs. var models for stock returns." *Unpublished paper, Chicago GSB* .

Conrad, J. & G. Kaul (1998): "An anatomy of trading strategies." *The Review of Financial Studies* **11(3)**: pp. 489–519.

DITTMAR, R. F. (2002): "Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns." *The Journal of Finance* **57(1)**: pp. 369–403.

EASLEY, D. & M. O'HARA (1987): "Price, trade size, and information in securities markets." *Journal of Financial Economics* **19(1)**: pp. 69 – 90.

EASLEY, D. & M. O'HARA (1992): "Time and the process of security price adjustment." *The Journal of finance* **47(2)**: pp. 577–605.

ENGLE, R. F. (1982): "Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation." *Econometrica* **50(4)**: pp. 987–1007.

FAMA, E. F. (1991): "Efficient capital markets: Ii." *The Journal of Finance* **46(5)**: pp. 1575–1617.

FEDERAL RESERVE BANK OF ST LOUIS (2018): Available at `https://fred.stlouisfed.org/`.

FORESI, S. & F. PERACCHI (1995): "The conditional distribution of excess returns: An empirical analysis." *Journal of the American Statistical Association* **90(430)**: pp. 451–466.

GRANGER, C. W. & A. ANDERSEN (1978): "On the invertibility of time series models." *Stochastic processes and their applications* **8(1)**: pp. 87–92.

HARVEY, C. R. & A. SIDDIQUE (2000): "Conditional skewness in asset pricing tests." *The Journal of Finance* **55(3)**: pp. 1263–1295.

HASBROUCK, J. (1991): "Measuring the information content of stock trades." *The Journal of Finance* **46(1)**: pp. 179–207.

HORNIK, K. (1993): "Some new results on neural network approximation." *Neural Networks* **6(8)**: pp. 1069 – 1072.

JEGADEESH, N. & S. TITMAN (1993): "Returns to buying winners and selling losers: Implications for stock market efficiency." *The Journal of Finance* **48(1)**: pp. 65–91.

KAHNEMAN, D. & A. TVERSKY (1982): "The psychology of preferences." *Scientific American* **246(1)**: pp. 160–173.

Koenker, R. & G. Bassett Jr (1978): "Regression quantiles." *Econometrica: journal of the Econometric Society* pp. 33–50.

Koenker, R. & K. F. Hallock (2001): "Quantile regression." *Journal of economic perspectives* **15(4)**: pp. 143–156.

Koenker, R. W. & V. d'Orey (1987): "Algorithm as 229: Computing regression quantiles." *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **36(3)**: pp. 383–393.

Park, C.-H. & S. H. Irwin (2007): "What do we know about the profitability of technical analysis?" *Journal of Economic Surveys* **21(4)**: pp. 786–826.

Pedersen, T. Q. (2015): "Predictable return distributions." *Journal of Forecasting* **34(2)**: pp. 114–132. For.2323.

R Core Team (2013): *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rapach, D. & G. Zhou (2013): "Chapter 6 - forecasting stock returns." In G. Elliott & A. Timmermann (editors), "Handbook of Economic Forecasting," volume 2 of *Handbook of Economic Forecasting*, pp. 328 – 383. Elsevier.

RStudio Team (2015): *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA.

Sharpe, W. F. (1964): "Capital asset prices: A theory of market equilibrium under conditions of risk." *The Journal of Finance* **19(3)**: pp. 425–442.

Timmermann, A. & C. W. Granger (2004): "Efficient market hypothesis and forecasting." *International Journal of forecasting* **20(1)**: pp. 15–27.

Trippi, R. R. & D. DeSieno (1992): "Trading equity index futures with a neural network." *Journal of Portfolio Management* **19**: pp. 27–27.

Welch, I. & A. Goyal (2008): "A comprehensive look at the empirical performance of equity premium prediction." *The Review of Financial Studies* **21(4)**: pp. 1455–1508.

Yang, J. W. & J. Parwada (2012): "Predicting stock price movements: an ordered probit analysis on the australian securities exchange." *Quantitative Finance* **12(5)**: pp. 791–804.

Appendix A

Autocorrelation plots

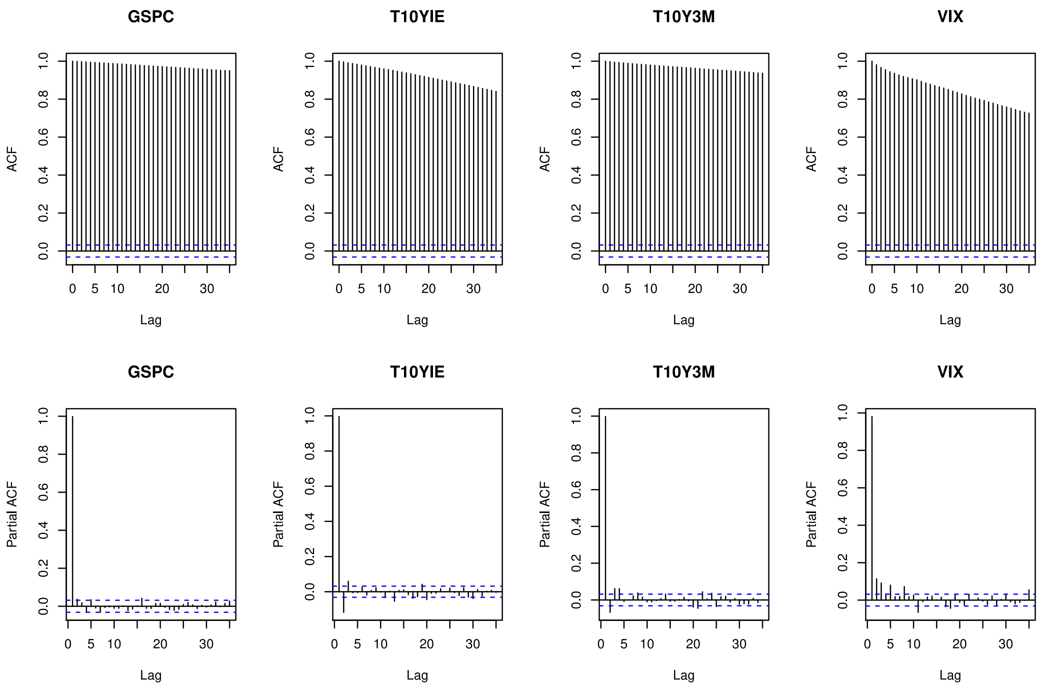

Figure A.1: Auto-correlation and partial auto-correlation functions of the data

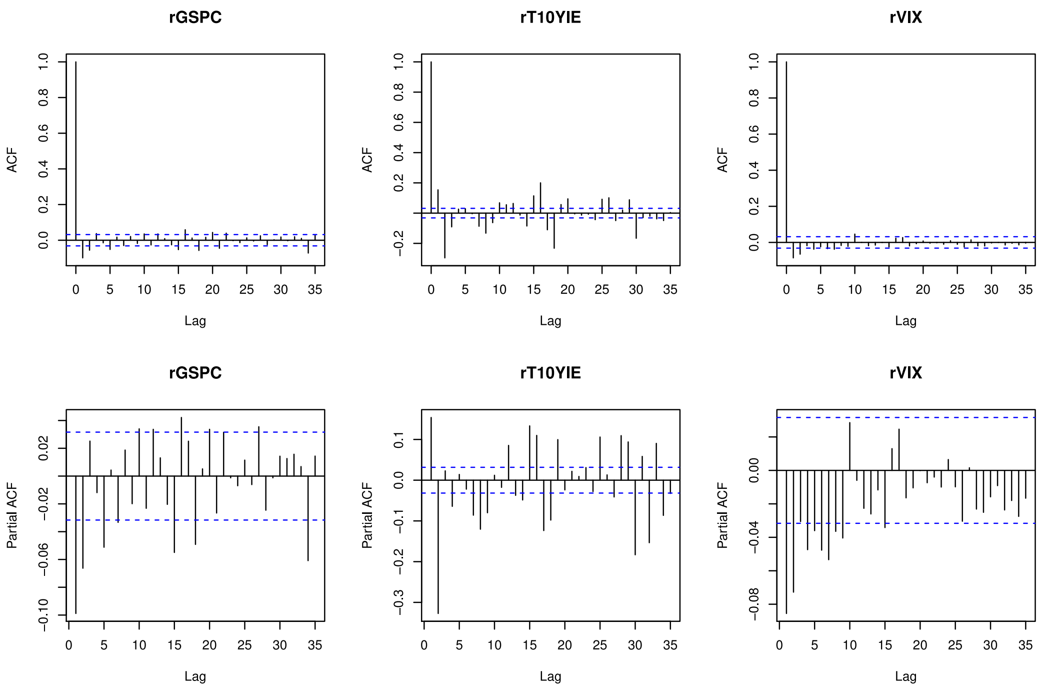

Figure A.2: Auto-correlation and partial auto-correlation functions of the returns on data

Appendix B

Content of Enclosed DVD

There is a DVD enclosed to this thesis which contains empirical data in csv format and R source codes in plain text format.

- Folder 01: r_source_codes
- Folder 02: data

R software and Rstudio as a free and open-source integrated development environment for R and supplementary packages available on the official R project website were used for all computations, graphical and tabular outputs. The data were fetched by the help of function `getSymbols` from the `quantmod` package. All graphs were created with `ggplot2` and combined by `ggarrange` from the package `qqpubr`. To create LaTeX format tables the packages `xtable` and `stargazer` served very well. The quantile regressions estimates were performed by the function `rq` from the package `quantreg`. The GARCH models were estimated by the help of `fGarch` and `rugarch` packages.