Applications of Quantiles In Portfolio Management

Fredrik Bergling

Applications of Quantiles In Portfolio Management

Fredrik Bergling

Advisor: Maciej Klimek

Uppsala University, Uppsala

Master Degree project 30 Credits

2018-06-18

Contents

Abstract

The purpose of this study is to look at how we use quantiles within portfolio management. We have a look at a couple of well known quantile based concepts from an mathematical perspective. For example Value at Risk and Conditional Value at Risk. We also look at how we can use quantile regression to get a more robust prediction of expected return. The most important conclusion that we come up with is that we can see an advantage when doing portfolio management by obtaining expected return using quantile regression.

1   Introduction

The world of Portfolio Management we have, for a long time, has relied on the ideas of Markowitz (1952), where the relationship between expected return and standard deviation to perform investment decision is explored. This has had an enormous impact on Portfolio Management and Asset Management. But this idea was based on averaging everything, this means that expected return of an asset is calculated by averaging the historical returns of a certain time period, which creates problems as pointed out by Savage (2009). Later, Kaplan (2011) presented an improved version of Markowitz's portfolio theory where he for example substitutes the standard deviation for Conditional Value at Risk and looks at the tail risk instead of the average variation. This a very interesting idea since when minimizing the standard deviation, this effects large negative returns but also large positive returns. Instead, by using the Conditional Value at Risk, you can simply look at the largest 5% negative returns and focus on minimizing these.

A big part of portfolio theory is the question regarding the expected return, how to get the best prediction of future return and thereby create the best portfolios.To our help here, there are a lot of different asset pricing models. Some important models within asset pricing and asset allocation are the basic CAPM and Fama-French models, which both are regression based where we usually use the Ordinary Least Square method for estimating beta. This could create a problem of missing the effect of extreme events. In this paper I will examine how such a problem could be eliminated by doing the regression over different parts of the sample with quantile based regression methods to capture the effect of outliers and get a better strategy for Portfolio Management.

In the last 20 years we had two major crises in the stock markets around the world. The first one was around 2001, and was an effect of the IT bubble burst. It burst after years of drastically increasing stock prices, in particular in the IT stocks. During the IT crash some markets experienced decrease up towards 70%. The Dow Jones Industrial Average, that we are evaluating in this paper didn't collapse as badly as many other markets since it consisted mostly of big steady industrial companies. The other crisis worth mentioning is the financial crisis of 2008, where Lehman Brothers collapsed and went bankrupt. During this crisis the Dow Jones did take a harder hit

than it did in 2001 and we saw a decline of around 50%. This crisis had big consequences for both countries and the financial industry, where some countries were forced to step in and save their banks for the sake of the future of the financial industry. Both of these crises have been followed by major increases in stock prices. So if you could minimize the tail of the returns you could get a far better portfolio return. Therefore, this paper will explore how we could use quantiles in portfolio theory to get a better return over time.

2 Quantiles

2.1 Background/Theory

Quantiles are commonly used in statistics and probability theory.

Definition of quantiles according to Acerbi, Tasche(2002):

$x_\alpha = q_\alpha(X) = \inf\{x \in R : P[X \le x] \ge \alpha\}$ is the lower α -quantile of X

$x^\alpha = q^\alpha(X) = \inf\{x \in R : P[X \le x] > \alpha\}$ is the upper α- quantile of X

A quantile is the cut off points in a distribution, that makes all the samples have the same probabilities. There always is one less cut off point than there are number of groups. The 50th quantile is the median, the 25th and 75th quantiles are called quartiles and the 100th is called the percentile.

Quantiles is a very well used concept within statistics, you can see in news papers or magazines. For example, it is probable that most of us have read a newspaper article about how much the richest one percent is earning, this is an example of the 100-quantile of the distribution of income with the cut off at the 99th percentile. Quantiles can also be used a lot within econometrics, for example when locking at income difference in a country, we can use quantiles to separate different groups of the labor force. This can also be interesting when grading papers, to see how a certain quantile of the students is doing.

If X is a real-valued random variable and F_X denotes the cumulative distribution function for X that is,

$$F_X(x) = \int_{-\infty}^{x} f_X(z)dz, \tag{1}$$

where x ∈ R and $f_X(x)$ is the probability density function, provided that f_X exists.

We then define the corresponding quantile function as the generalized inverse of F:

$$F_X^{-1}(q) = \inf\{x : q \leq F_X(x)\}, \qquad q \in (0,1) \qquad (2)$$

Then $F_X^{-1}(q)$ is the lower q-quantile of X.

Then we can say that the lower q-quantile x_q is the lowest value of x so that the probability of X not being bigger than x is not smaller than q.

As mentioned above the most used regression method in finance and econometrics is linear regression and one of the most effective linear regressions is the OLS method. Ordinary Least Square calculates the coefficients and a constant that describes the mean relation between a dependent and an independent variable. It measures the average change in the dependent variable when you change the independent variable with one unit.

In this paper we are presenting an alternative regression method, quantile regression. The idea of quantile regression is in some sense the similar to the OLS regression. In quantile regression, we look at the tendency for a certain quantile instead of the central tendency as we do in OLS. Quantile regression is preferable to OLS in some cases, because it is better at capturing the effect of outliers and it can look at different parts of the sample. It is also a more robust method than the OLS. Both of these methods can be presented as a optimization problem.

If we recall how to get the unconditional mean, we could present this as a optimization problem. The idea is that we need to find the μ that minimizes the sum of squared residuals. Then μ will be the unconditional mean of the vector y and the optimization problem that needs to be solved is the one below:

$$\arg\min_{\mu \in R} \sum (y_i - \mu)^2 \qquad (3)$$

If we now want to approximate the conditional mean by a linear function

$E(Y|X = x) = x\beta + \alpha$, we have to solve a similar optimization problem as the previous one but we replace μ with $\beta x_i + \alpha$:

$$\arg\min_{\beta,\alpha \in R} \sum (y_i - \beta x_i - \alpha)^2 \tag{4}$$

When we are optimizing (4), we are looking for the central tendency between Y and X. The β that we are solving for, explains what happens on average to y when we have a change in x and α accounts for the constant.

Now we are moving on to the part that gives us the quantile regression. A quantile can be seen as the cut off point in a sample. It can be found by minimizing the weighted absolute sum of residuals by varying μ. In the formula below we will find the median since it is symmetrically weighted, but when we later introduce asymmetrical weighting, this allows us to be able to do the same thing to get any quantile.

So the optimization problem that gives the median of a sample is as below:

$$\arg\min_{\mu \in R} \sum |y_i - \mu| \tag{5}$$

We can show that this works through an example, we assume a sample of 5 arbitrary numbers $\{s_1, ..., s_5\}$ in a plot created by the output of formula above. In the example below s=$\{1, 2, 3, 4, 5\}$, This can be shown for any number and all sizes of the sample, but for simplicity have I stayed at 5 numbers.

$$\sum_{i=1}^{5} |s_i - \mu| \tag{6}$$

When varying μ we come up with the following plot, where it is clear that that minimum is at 3 which also is the median.

Sum of absolute least value
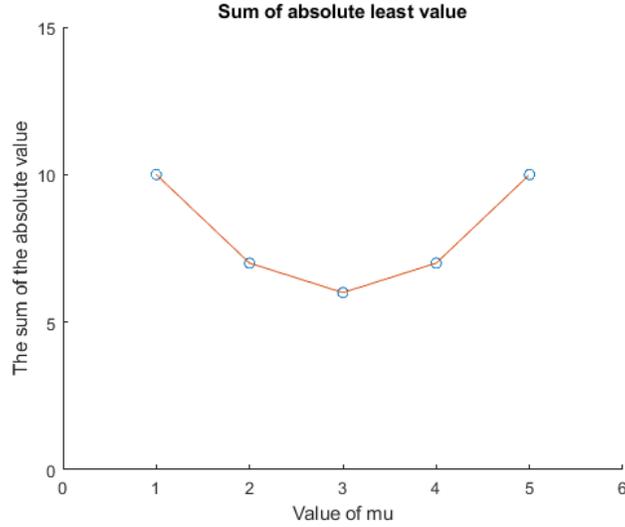
(a)

Using the same idea as when we created the linear approximation of the conditional mean function, we now create the linear conditional median function. We just substitute μ with $\beta x_i + \alpha$ in the optimization problem as below.

$$\underset{\beta,\ \alpha \in R}{\arg\min} \sum |y_i - \beta x_i - \alpha| \tag{7}$$

We encounter the same problem here as in the case of the conditional mean function, that if the sample is not standardized, we will not find the correct beta and we need to introduce a constant as below.

When we are doing median regression, we are minimizing symmetrical weighted sum of the absolute residuals and when we are doing regression for the other quantiles, we will use a asymmetrical weight of the sum of the absolute residuals. To achieve this asymmetrical weight we introduce the tilted absolute value function ρ_τ.

$$\rho(z) = z(\tau - I_{z<0}) \tag{8}$$

Below, we can see a figure of how the function ρ could look like.

(a)
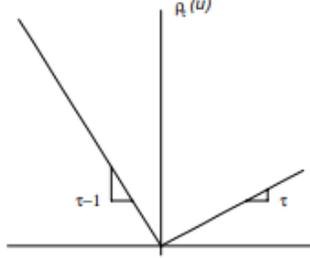
In the figure above we see a graph over the tilted value function. (Allen, Powell and Singh, 2011)

The function ρ_τ depends on τ, which is the variable that decides which quantile we are doing the regression on. The parameter τ can have a value between 0 and 1, whereas $I_{z<0}$ which is the indicator function that will take value 1 if z is smaller than 0 and 0 if z is bigger than 0. This way ρ is punishing the samples that give negative residuals with a negative weight. When $\tau = 1/2$ we have the symmetrical weight and the regression will be a median regression as shown above. The proof of this will be shown here below:

We can start by letting X be a random variable with CDF F_X. Now we will prove that we get a certain quantile from the formula below:

$$F_X^{-1}(q) = \arg \min_z E[\rho_q(X - z)] \tag{9}$$

We now want to minimize the differentiable function,

$$h : R \ni z \mapsto (q - 1) \int_{-\infty}^{z} (x - z)dF_X(x) + q \int_{z}^{\infty} (x - z)dF_X(x) \in R \tag{10}$$

By the Leibnitz Integral Rule,

$$h'(z) = -(q - 1) \int_{-\infty}^{z} dF_X(x) - q \int_{z}^{\infty} dF_X(x) \tag{11}$$

9

If we then re-arrange this when keeping in mind that $F_X(\infty) = 1$ and $F_X(-\infty) = 0$, we get the following:

$$h'(z) = F_X(z) - q \tag{12}$$

since we know that the CDF is non-decreasing, we also know that we have a minimum at the point z

So now we end up with this formula below when replacing the absolute value by $\rho(\cdot)$:

$$\arg\min_{\beta,\alpha \in R} \sum \rho_\tau(y_i - \beta_\tau x_i - \alpha) \tag{13}$$

Since $\beta_\tau x_i + alpha$ is a linear function this can be solved by using linear programming methods. Preferably the Simplex method, which is a method in mathematical optimization, developed by Gerorge Datzig in 1946. When we solved this we got the β and α for the qth-quantile regression.

Chan and Lakonishok (1992) proposed that, to get the best result, we need to calculate several βs and αs for different quantiles and then take a weighted average of these to get a more robust β and α than the one you receive from the Ordinary Least Square. In this paper I will use two different weighted averages, that is both of the weighted averages that Allen, Powell and Singh(2011) proposed. The ones which I will be using are Tukey's trimean and a symmetrical weight.

Tukey's trimean was first introduced by Arthur Bowley. But it got it's breakthrough when it was mentioned by John Tukey (1977). The idea of Tukey's trimean is that you take a weighted average of the median and the two quartiles. As the formula below

$$TM = \frac{Q_1 + 2Q_2 + Q_3}{4} = 0.25Q_1 + 0.5Q_2 + 0.25Q_3$$

In our case, we take one quarter of the βs from the quantile regression of the 25th quantile and 75th quantile and then take half of the median regression β and then sum these as below:

$$\beta_t = 0.25\beta_{0.25,t} + 0.5\beta_{0.5,t} + 0.25\beta_{0.75,t}$$

This will give a more robust β for the regression. The other method to combine these βs into one coefficient is by symmetrical weight. This is also

a weighted average, but in this way it also captures the behavior of the 5th and 95th percentile by taking away some of the effect from the quartiles. The formula for this is presented below:

$$\beta_t = 0.05\beta_{0.05,t} + 0.2\beta_{0.25,t} + 0.5\beta_{0.5,t} + 0.2\beta_{0.75,t} + 0.05\beta_{0.95,t}$$

This will give a greater weight to the more rare events than the Ordinary Least Square method and that way give a more robust explanation of the returns over all of the quantiles and we are doing the same thing for the αs to get a more robust intercepts as well.

2.1.1 Visualization

To get a better overview of quantiles we can use a box plots. It is mostly used in descriptive statistics. The idea of the plot is that we have a rectangle and two whiskers which are T-shaped. One is attached to the top of the rectangle and the top of this T-shape marks the largest value of this sample or the limit large outliers. The other whisker is turned upside down and attached to the bottom of the rectangle. The bottom of the upside down whisker shows the lowest value of the sample or limit for small outliers. Then we have the rectangle which contains half of the sample. The bottom of the rectangle marks the lower quartile and the top of the rectangle marks the upper quartile. Then there is usually a line on the rectangle which marks the median.

The Interquartile Range or IQR is the difference between the third and first quartiles. It is used to create a graphical representation of a probability distribution, and it can also be used to grade outliers. An outlier is said to be mild if it is more than 1.5 IQR away from the first or the third quantile but not more than 3 IQR away. If it is more than 3 IQR away from the first or the third quantile, the outlier is called extreme.

2.2 Application in Portfolio Management

The applications for quantiles and quantile regression are quite many within portfolio management and in finance. Many of the most famous asset pricing models are regression based, for example CAPM, the single-factor model and

the multi-factor model.

If we begin by looking at the basic ideas of Portfolio Management, which is to get the highest expected excess return when bearing a certain level of risk. One of the most established measures of how good a portfolio is, is the Sharpe ratio, which is the ratio between the excess return and standard deviation as in the formula below.

$$\frac{\bar{R}_p - R_f}{\sigma_p}, \tag{14}$$

where
$\bar{R}_p$ is the expected return of the portfolio,
R_f is risk-free rate,
σ_p is the standard deviation of the portfolio.

In finance, we have a risk measurement that calculates the average loss with a given level of probability and that is Conditional Value at Risk. It states how much an investment can loose over specific time period with a given probability. But we start with the more common Value at Risk, which not is a coherent risk measure. Value at Risk, VaR in short, is a quantile based method.

The definition of Value at Risk is as follows, when letting X be a representation of a random variable of the gain(with loss being negative):

$$\mathrm{VaR}_{(1-\alpha)}(X) = \inf\{c \in R : P(-X \le c) \ge 1 - \alpha\} = F_{-X}^{-1}(1 - \alpha)$$

Where α is a chosen tolerance level, that takes a value $0 < \alpha < 1$.

The definition of a coherent risk measure is:

- non-decreasing $(X \le Y \Rightarrow v(X) \le v(Y))$

- translation invariant $(v(\lambda + X) = v(X) - \lambda$ for $\lambda \in R)$

- positive-homogeneous $(v(\lambda X) = \lambda v(X)$ for $\lambda \ge 0)$

- subadditive $(v(X + Y) \le v(X) + v(Y))$

The last two are sometimes replaced with weaker requirements of convexity.

If we want to look at an example if Value at Risk is a coherent risk measure, we can assume that there exists two possible investments, Invest 1 and Invest 2. The loss distribution for these investments you can see in the figure below.

Investment choice	Scenario1	Scenario 2	Scenario 3
Probability of occurrence	0.03	0.03	0.94
Invest 1	-500	0	0
Invest 2	0	-500	0

If we now use this to calculate the VaR at a 5% tolerance level for Invest 1, Invest 2 and Invest 1 + Invest 2. This gives us

$$\text{VaR}_5(\text{Invest 1}) = \inf\{c \in R : P(\text{Invest1} \leq c) \geq 0.95\} = 0$$

$$\text{VaR}_5(\text{Invest 2}) = \inf\{c \in R : P(\text{Invest2} \leq c) \geq 0.95\} = 0$$

$$\text{VaR}_5(\text{Invest 1+ Invest 2}) = \inf\{c \in R : P(\text{Invest 1 + Invest 2} \leq c) \geq 0.95\} = -500$$

Since $\text{VaR}_5(A + B) \leq \text{VaR}_5(A) + \text{VaR}_5(B)$ is violated in the example above, this means that VaR is not subadditive. So this makes Value at Risk a non-coherent risk measure, but the Conditional Value at Risk that we will define below, is a coherent risk measure and that is why it has become more and more popular. It is included as a risk measure in the improvement of Markowitz's portfolio theory by Kaplan (2011).

The definition of Conditional Value at Risk according to Acerbi & Tasche(2002) assume that $E[X^-] < \infty$. Then

$$\text{CVaR}_\alpha(X) = \inf\{(E[(X - s)^-])/\alpha) - s : s \in R\}$$

is the CVaR at tolerance level α of X.

This gives you the Conditional Value at risk or expected shortfall, which is a different name for the same thing. It can be interpreted as the mean of the loss in the quantile between α and zero.

The definition of Expected shortfall according Acerbi, Tasche(2002).

$$ES_\alpha(X) = -\alpha^{-1} \int_0^\alpha q_u(X)du.$$

For example: if we are assuming the following profit distribution at some time in the future it looks as follows:

Probability	Profit
1 %	-100
4 %	-50
10 %	-25
25 %	-10
40 %	0
15 %	50
5 %	100

If we have a look at the table above we could easily see the VaR value. The profit at the tolerance level 1% is -100 and at 5% is -50, this means that with 95% probability you will not lose more than 50.

q	expected shortfall
1%	-100
5%	-60
10%	-42.5
25%	-26

3   Regression based asset pricing models

We can see in the table above that Conditional Value at Risk answers another question when compared to Value at Risk. The answer that Conditional Value at Risk gives is that the worst 5% case scenario gives expected loss of 60. Instead of the answer that Value at Risk gives which is that with 95% probability you won't lose more than 50.

As mentioned above there are plenty of regression based models for determining future returns on assets. For example CAPM, which an abbreviation of Capital Asset Pricing Model. The formula for CAPM is as follows:

$$E[R_i] = \bar{R}_i = R_f + \beta_i(E(R_m) - R_f)$$

The idea is that all stocks have some correlation with the market risk premium. So in this model you just have to estimate the expected return of the market. Then the r_f are known and β_i will be evaluated with help of a linear regression between historical return of the asset and the market. This regression could be done in many different ways. The most used method is OLS. But you could also do this with a combination of quantile regressions that are weighted with some kind of weighted average like they did in Allen, Powell and Singh (2011) as we are trying to do in this paper. In Allen, Powell and Singh (2011) the authors are trying to use βs generated by quantile methods in the Fama-French Three-factor model to create a portfolio with a higher Sharpe ratio in times of financial distress. They succeed creating a risk premium compared to the standard OLS.

The Fama-French three-factor model is a widely used asset pricing model, it is an evolution from the classic CAPM model. It was designed by Eugene Fama and Kenneth French. The main idea of CAPM is as mentioned before in this paper, that all stocks have a certain correlation with the market portfolio and therefore you don't have to evaluate the expected return for each asset. Instead you can find the correlation between the market and the assets and when you know this you can just evaluate the expected return for the market to get the expected return of the asset by multiplying these two together. However Fama and French thought that this didn't explain enough of the asset's return so the they introduced two more factors, coming up with the model below.

$$E[R_i] = R_f + \beta_i \text{avg}(R_m - R_f) + \beta_{\text{SMB, i}} \text{avg}(SMB) + \beta_{\text{HML,i}} \text{avg}(HML) + \alpha$$

As we see in the model above the first part of the model looks exactly like CAPM but then we have the other two added factors. The first one is the SMB factor, which is the abbreviation of Small minus Big. This factor is the difference between the return of small companies and return of big companies. Fama and French's idea is that smaller companies have a higher return due to a higher risk premium that smaller companies require. The second factor that is added is HML, which is the abbreviation of High minus Low, this refers to the book to market ratio. You can also see this as the difference between value stocks and growths stocks. Value stocks are big companies that probably not are going to grow that much but are paying a great

dividends every year and are steady. While most growth stocks are newly started companies or at least smaller ones with a great idea that haven't started earning that much money yet, but are growing. Now when we have these two additional factors we do a linear regression on historical data with these factors as independent variables and the assets historical return to get the betas for each of the factors and assets. This linear regression is usually done by Ordinary Least Square as mentioned before but in this paper we are doing it by Quantile Regression.

There are also benefits from this type of multi-factor models when calculating the standard deviation. Instead of calculating the standard deviation for each of the assets, we can just calculate it for the factors and then just multiply it with the corresponding β as in the formula below.

$$\sigma_i^2 = \beta_i \text{var}(R_m - R_f) + \beta_{\text{SMB, i}} \text{var}(SMB) + \beta_{\text{HML,i}} \text{var}(HML)$$

Then when entering into a portfolio environment the expected return formula is quite straightforward. It is just the sum of all the expected returns, but here you can see one of the benefits of multi-factor models. Instead of trying to forecast n number of returns, we just have to forecast three factor returns and then we can just multiply them with the coefficient to get an expected return for the portfolio.

$$E[R_p] = R_f + \sum_{i=1}^{n} \beta_i \text{avg}(R_m - R_f) + \beta_{\text{SMB, i}} \text{avg}(SMB) + \beta_{\text{HML,i}} \text{avg}(HML) + \alpha_i$$

Now when we are moving on to how we calculate the risk of the portfolio we also see that we have an advantage in the calculation effort. The first part of the σ_p formula is just the sum of the σ for the assets and the other part of this formula is the covariance. In the covariance part of the σ_p formula, we can see that we just have to multiply the variance of the factor with corresponding βs for two separate assets to get the covariance between these two assets.

$$\sigma_p^2 = \sum_{i=1}^{n} \beta_i \text{var}(R_m - R_f) + \beta_{\text{SMB, i}} \text{var}(SMB) + \beta_{\text{HML,i}} \text{var}(HML) +$$

$$\sum_{i=1}^{n} \sum_{i \neq j}^{n} \beta_i \beta_j \text{var}(R_m - R_f) + \beta_{\text{SMB, i}} \beta_{\text{SMB, j}} \text{var}(SMB) + \beta_{\text{HML,i}} \beta_{\text{HML,j}} \text{var}(HML)$$

When we have added the SMB and the HML to CAPM, the Fama-french three-factor model can explain over 90% of the return for a well diverisified portfolio, while CAPM is just explaining around 70%. So we can see are rather drastic improvement in rate of explanation. We can also see a drastic reduction in computational power when we compare it to the mean-variance portfolios.

4   Portfolio Theory

The modern Portfolio Theory is a very young science, it was first published by Harry Markowitz in 1952. This paper was later awarded the Nobel Prize. The idea that Markowitz came up with in this paper was the mean-variance portfolio optimization.

When we have the expected return for each asset, we get the expected return for the portfolio by multiplying the expected return of the asset with the corresponding weights of the asset.

$$E[R_P] = \bar{R}_P = \sum_i w_i E[R_i]$$

And then Markowitz proposed that the risk for the portfolio should be calculated by the formula below:

$$\sigma_P^2 = \sum_i w_i^2 \sigma_i^2 + \sum_i \sum_{j \neq i} w_i w_j \sigma_i \sigma_j \rho_{ij}$$

If we are looking at these calculations by vector algebra, we start with defining the different vectors. The first one is the weight vector of the portfolio.

$$w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$

The next vector in question is the expected return vector.

$$R = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{bmatrix}$$

and then we have the covariance matrix.

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{1,2} & \cdots & \sigma_{1,j} \\ \sigma_{2,1} & \sigma_2^2 & & \\ \vdots & & \ddots & \\ \sigma_{i,1} & & & \sigma_i^2 \end{bmatrix}$$

and when we have calculated the expected return in vector form, we will transpose the weight vector and multiply it with the expected return vector to get the expected return.

$$w^T \bar{R} = \bar{R}_p$$

And to get the risk for the portfolio we transpose the weight vector and multiply it with the covariance matrix, then we multiply this with the weight vector again as the formula below:

$$w^T \Sigma w = \sigma_p^2$$

Now we will discuss the Sharpe ratio for a while since this has a great impact onl in portfolio optimization and finance. The Sharpe ratio was developed by William F. Sharpe in 1966. It calculates the reward to risk ratio and is a good evaluation tool for portfolios as well as great formula to optimize the portfolio with.

$$\frac{R_p - R_f}{\sigma_p}$$

So now when we have the formula for the portfolio expected return and the portfolio standard deviation, we can formulate a portfolio optimization problem.

Portfolio optimization problem:

$$\max \frac{R_p - R_f}{\sigma_p}$$

With constraint:

$$\sum_{i=1}^{n} w_i = 1$$

$$w_i \geq 0, \text{for all i}$$

Which will give us the tangency portfolio with only positive weights, which is used in the case where short-selling is not allowed.
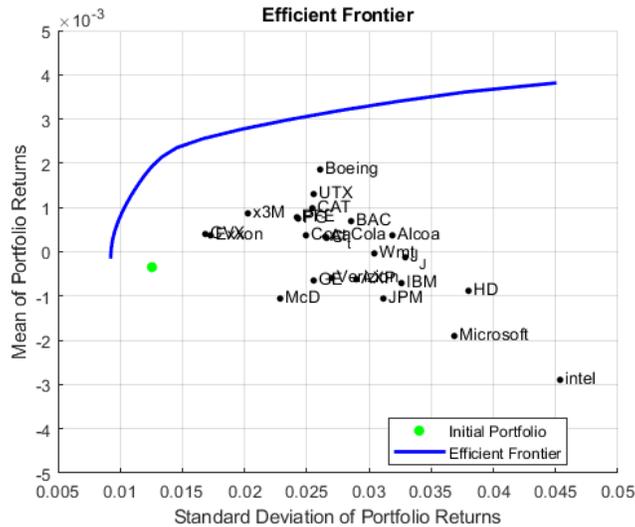Portfolio optimization problem when short selling is allowed:

$$\frac{R_p - R_f}{\sigma_p}$$

Subject to:

$$\sum_{i=1}^{n} w_i = 1$$

4.0.1   Efficient Frontier

The Efficient Frontier is a hyperbola of the portfolios with the highest expected return for each level of risk. It contains an endless number of portfolios.

(a)

In the figure above you can see the power of diversification, each of the points in the figure represents an asset in position of it's corresponding level of risk and expected return. So we can see that all the options of portfolios on the efficient frontier have a higher excepted return for the same level of risk.

One question that arises is which portfolio to choose? But since we have a diminishing derivative or negative second derivative, this means that we get a smaller increase in expected return the more we increase our level of risk. That is why the Sharpe ratio as we mentioned above is a useful tool to find the optimal portfolio. The portfolio with the highest Sharpe ratio is also called the Tangency Portfolio. This portfolio can be found by drawing a tangent from the risk-free rate on to the expected return axis, then it will touch the tangency portfolio. This is also the optimal portfolio with the highest Sharpe ratio. The tangent is called the Capital Allocation Line.

4.0.2 Global Minimum Variance Portfolio

On the efficient frontier we have an endless number of different portfolio options to choose from depending on your risk preferences. As mentioned above we have the tangency portfolio which is the portfolio with the highest expected return per unit of risk but there is also an other option which is very interesting, the Global Minimum Variance portfolio in case a risk-free rate doesn't exist. Here, the whole target is to minimize the the risk of the portfolio by the optimization problem below:

$$\min \frac{1}{2} \sum_{i,j=1}^{n} w_i w_j \sigma_{ij}$$

Subject to:

$$\sum_{i=1}^{n} w_i \bar{R}_i = \bar{R}$$

$$\sum_{i=1}^{n} w_i = 1$$

This portfolio is found farthest to the left on the Efficient Frontier, where the efficient meets the inefficient frontier.

We can solve this problem using Lagrange multipliers λ and μ. Then we will express the Lagrangian as follows:

$$L = \frac{1}{2} \sum_{i,j=1}^{n} w_i w_j \sigma_{ij} - \lambda(\sum_{i=1}^{n} w_i \bar{R}_i - \bar{R}) - \mu(\sum_{i=1}^{n} w_i - 1)$$

Now by taking the derivative for each of the weights and then setting it equal to zero we can solve this and get the Global Minimum Variance Portfolio.

We can look at this in a case with two variables, then the Lagrangian looks as follows:

$$L = \frac{1}{2}(w_1^2 \sigma_1^2 + w_1 w_2 \sigma_{12} + w_1 w_2 \sigma_{21} + w_2^2 \sigma_2^2)$$
$$-\lambda(\bar{R}_1 w_1 + \bar{R}_2 w_2 - \bar{R}) - \mu(w_1 + w_2 - 1)$$

Then by taking the derivative with respect to the unknown we get:

$$\frac{\partial L}{\partial w_1} = \frac{1}{2}(2\sigma_1^2 w_1 + \sigma_{12} w_2 + \sigma_{21} w_2) - \lambda \bar{R}_1 - \mu$$

$$\frac{\partial L}{\partial w_2} = \frac{1}{2}(\sigma_{12}^2 w_1 + \sigma_{21} w_1 + 2\sigma_2^2 w_2) - \lambda \bar{R}_2 - \mu$$

$$\frac{\partial L}{\partial \lambda} = \bar{R}_1 w_1 + \bar{R}_2 w_2 - \bar{R}$$

$$\frac{\partial L}{\partial \mu} = w_1 + w_2 - 1$$

now we can use the that $\sigma_{12} = \sigma_{21}$, when we set the derivatives equal to zero:

$$\sigma_1^2 w_1 + \sigma_{12} w_2 - \lambda \bar{R}_1 - \mu = 0$$
$$\sigma_{21} w_1 + \sigma_2^2 w_2 - \lambda \bar{R}_2 - \mu = 0$$
$$\bar{R}_1 w_1 + \bar{R}_2 w_2 = \bar{R}$$
$$w_1 + w_2 = 1$$

by solving this system of equations we will get the weights for the Global Minimum Weight. "The case of two assets is actually degenerate because the two unknowns w_1 and w_2 are uniquely determined by the two constrains"

(Luenberger, 1998, p.159). These type of calculations are only possible when we don't have any restriction about no short-selling.

If there exists a risk-free rate, the Global Minimum Variance portfolio won't be on the Efficient Frontier since there will be a portfolio that is a combination between risk-free rate and the Tangency portfolio that gives a higher expected return for the same level of risk. This line is as mentioned above called the Capital Allocation Line, but in the case where the portfolio only is allowed to contain risky assets. The basics behind this portfolio is that we like to find low variance assets or assets with low correlation to get to get the lowest possible variance for the portfolio.

4.0.3   Conditional Value at Risk Portfolio Optimization

Now we are moving on to Conditional VaR portfolio optimization, this an extension of modern portfolio theory and it is more commonly used nowadays. This is due to the fact that the Mean-variance optimization issues concerning the development of the covariance matrix and deciding the correlation between assets.

(a)
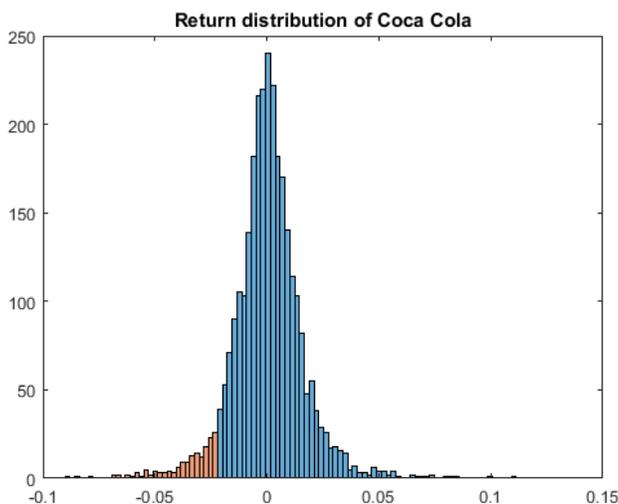
In the figure above we can see the distribution of the return for an asset and in the red columns, we have the 5% biggest loses. The idea of Condi-

tional Value at Risk as mentioned before, is to take the average of these red columns to get an idea of how much the portfolio's expected loss is in case of the worst 5%-scenario occurs.

The optimization problem that you need to solve to get the optimal weight for the CVaR portfolio is the one below.

$$\min_{w,\xi} \sum_{i=1}^{n} -E[y_i]w_i \tag{15}$$

subject to

$$\xi + (1-\alpha)^{-1} \sum_{j=1}^{J} \pi_j z_i \leq \omega \sum_{k=1}^{n} q_k w_k^0 \tag{16}$$

$$z_j \geq \sum_{i=1}^{n} (-y_{ij}w_i + q_i w_i^0) - \xi, \quad z_j \geq 0, \quad j = 1, ..., J \tag{17}$$

$$q_i x_i \leq \nu_i \sum_{k=1}^{n} q_k w_k, \quad i = 1, ..., n, \tag{18}$$

$$w_i - w_i^0 = u_i^+ - u_i^-, \quad i = 1, ..., n \tag{19}$$

$$0 \leq u_i^- \leq \bar{u}_i^-, \quad 0 \leq u_i^+ \leq \bar{u}_i^+ \quad i = 1, .., n \tag{20}$$

Where:
y_i is the scenario-dependent prices
w_i is the optimal weight
α is the specified probability level
π are the probabilities of scenarios y_j
z_j is a dummy variable
ω is a percentage of the initial portfolio value allowed for risk exposure.
q is the initial price of assets
ν is the maximum weight for one asset

By solving this optimization, we will find the optimal weights, the VaR which in the problem is ξ and the maximum return. By changing the risk ω

we get the CVaR efficient frontier.

The loss function over a certain period is:

$$f(w, y; w^0, q) = -y^T w + q^T w^0 \tag{21}$$

Since the loss function is convex in x so is the α-CVaR function and then the set of linear constraints, is 16 and 17.

Regarding 18, this is the value constraint. It is constructed to not give any asset a too big weight in the portfolio. This constraint is mostly used and makes the most sense when you are not allowing any short positions.

Equation 19 and 20 are regarding liquidity constraint. 19 is if a large transaction could change the price of the asset in question. 20 is if the positions are bound themselves. For a more thorough explanation of this optimization see Krokhmal, Palmquist and Uryasev (2001).

When we compare the Conditional Value at Risk optimization with the Mean Variance Portfolio optimization we will see that when the loss function is normally distributed, these two ideas will create the same Efficient Frontier. But when we have a non-normal distribution and especially non-symmetric distribution, we will see a significant difference in the frontiers from the CVaR and Mean-Variance approach. One very interesting difference between CVaR and MV is that CVaR looks at one of the tails of the return distribution, more precisely the high loss end of the distribution, and wants to reshape the tail of high losses. If we on other hand look at Mean-Variance we try to minimize the variance which is affecting both the gain and the loss tail. This means that it is looking at reshaping the high loss tail but it will also limit the high gains.

5   Data and Methodology

The data in this paper is contains of daily prices from 24 stocks, that are traded on the New York Stock Exchange and are included in Dow Jones Industrial Average. The period is between March 2000 to December 2017. We also have data from the same period for the Fama-French factors. These you can find on the Fama-French website, the web address is listed under

references. These are used to obtain the coefficients for the Fama-French factors. In the table below you can see the stocks included in this study.

3M	Alcoa	American Express
AT& T	Bank of America	Boeing
Caterpillar	Chevron	Citigroup
Coca Cola	Exxon	General Electrics
Home Depot	IBM	Intel
Johnson & Johnson	JP Morgan	McDonalds
Microsoft	Pfizer	Procter & Gamble
United Technologies	Verizion	Walmart Stores

The way that I am approaching the question is with quantile regression to get the coefficient for the Fama-French Three-Factor model. They will be calculated in five different quantiles. The quantiles will be 0.05, 0.25, 0.5, 0.75 and 0.95. Then these will be one single coefficient by applying Tukey's trimean and by using a symmetrical weighing. This is to study the behavior of the returns in different quantiles of the distribution. I will be using the financial tool box in MATLAB to create the optimal portfolio weights on the data from the previous year. I will try to find the max Sharpe ratio portfolio when we are looking at comparing with the mean variance optimization and when we are looking in the Conditional Value at Risk environment we will look at the portfolios with the smallest Conditional value at risk. When we have the portfolio weights, the hold out period is one year.

6   Creating Expected Return

As mentioned above we are using the Fama-French model to estimate the expected return. In this model we have to find the average for every factor and the regression coefficient associated with each company. Below I have presented all of the different steps in the process of obtaining the expected return from the quantile regression with symmetrical weighting.

The first step in the process is to execute the quantile regression. I do this for all the factors, the market risk premium, the SMB and the HML at 5%, 25%, 50%, 75% and 95%. We can see the result from this calculation below.

The betas for the market premium over the different quantiles

(a)
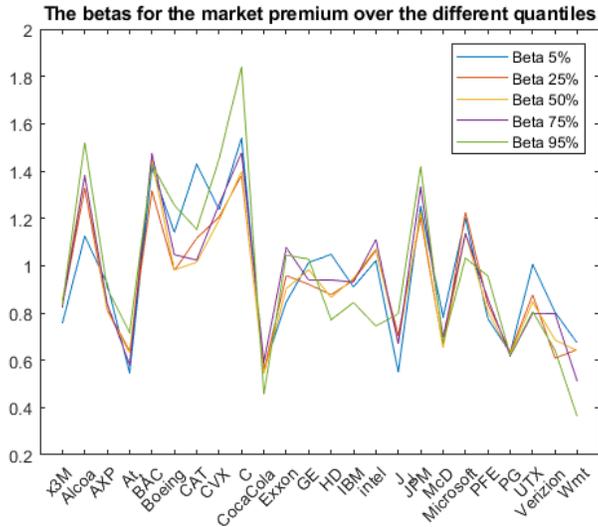
In the figure above we see the coefficient for the market premium for each company. We can see that the βs varies between 0.4 and 1.8. This is interesting since all companies have a positive relation with the market, which means that they are going in the same direction as the market over time. We can also see that both the extremes are developed by doing the quantile regression at 95%. This can be interpreted as that the relation between the market return and the stock return is changing in different market climates.

The next step is to create the β for the HML (High Minus Low) and below here we can see a figure of these for the different parts of the sample.

The betas for the HML over the different quantiles
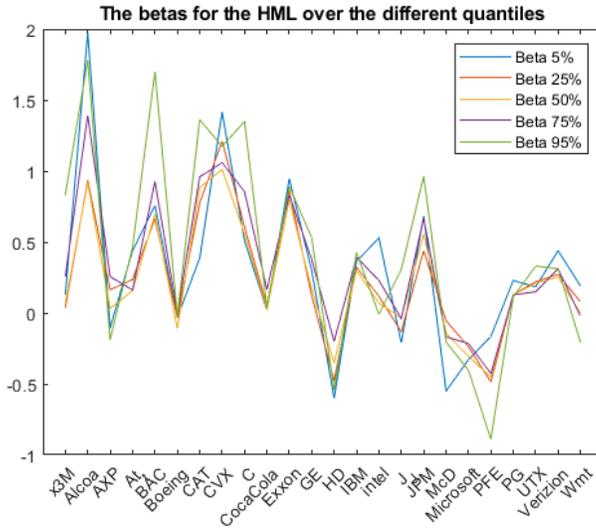
(a)

We can see that these βs are varying between -0.5 and 2. The interesting thing is that we have some companies that have a negative relation to this factor. The highest β is created by the regression at 5% and the lowest is created by the 95%. We can also see that there are significant differences between the β over the different quantiles.

Now we just have the SMB (Small Minus Big) left to look at.

(a)
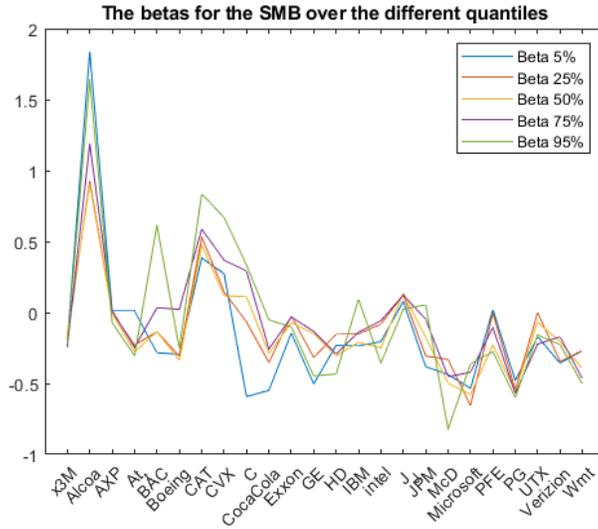
In the figure above we can see that the βs are varying between -0.8 and 1.8. But most of the βs are bound between -0.5 and 0.5. Here, we also have, as in the previous figure, a significant difference between the different quantiles.

Now when we have all the βs for the different factors and different quantiles, we will use the symmetrical weighted average to combine these into one β for each company and factor.

$$\beta_t = 0.05\beta_{0.05} + 0.2\beta_{0.25} + 0.5\beta_{0.5} + 0.2\beta_{0.75} + 0.05\beta_{0.95}$$

The βs we come up with are the ones in the table below.

Company	Mkt-rf	High Minus Low	Small Minus Big	Alpha
3M	0.8362	-0.1725	0.1526	2.9683e-04
Alcoa	1.3634	1.0345	1.1377	-0.0021
AXP	0.8332	0.0199	0.1174	-4.5497e-04
AT-T	0.6234	-0.2545	0.1987	6.6974e-04
BAC	1.4128	-0.0809	0.7506	-7.3815e-04
Boeing	1.0129	-0.2629	-0.0811	-1.8961e-04
CAT	1.0552	0.5310	0.8955	1.9987e-04
CVX	1.2131	0.2141	1.0707	-2.6507e-04
C	1.4424	0.0927	0.6553	-8.5335e-04
Coca-Cola	0.5535	-0.2954	0.0604	4.9834e-04
Exxon	0.9530	-0.0590	0.8282	-9.7801e-05
GE	0.9647	- 0.2118	0.2239	6.0703e-04
HD	0.8844	-0.2762	-0.3662	5.3673e-04
IBM	0.9326	-0.1705	0.3324	-7.3690e-05
Intel	1.0486	-0.1849	0.0928	2.6733e-04
JJ	0.6926	0.1191	-0.0884	2.2305e-04
JPM	1.2576	-0.1780	0.5883	-2.4481e-04
McD	0.6706	-0.4683	-0.1508	8.0491e-04
Microsoft	1.1445	-0.5361	-0.2658	4.5193e-04
PFE	0.8241	-0.1308	-0.4512	-3.8471e-04
PG	0.6388	-0.5353	0.1171	1.7229e-05
UTX	0.8491	-0.0945	0.2028	-4.7207e-04
Verizion	0.6982	-0.2362	0.2795	7.5959e-05
Walmart	0.5955	-0.3791	0.0054	-5.4353e-04

In the table above, we have the βs that we need to estimate expected return. Now we just have to find the expected return for each of the factors to estimate the expected return for each company.

7   Portfolio based on Variance

Now that we have created the expected return as we did in the previous section, we have to solve a optimization problem. This is effectively done using a linear programming method. The optimization problem that I have chosen to use to optimize is to maximize the Sharpe ratio, where we get the highest risk premium per unit of risk.

$$\max \frac{\bar{r}_p - r_f}{\sigma_p} \qquad (22)$$

Subject to:

$$\sum_{i=1}^{n} w_i = 1$$

$$w_i \geq 0$$

Where
$\bar{r}_p$ is the expected return of the portfolio
r_f is risk-free rate
σ_p is the standard deviation of the portfolio

The constraints that we are using are that no short-sale is allowed and the portfolio must be fully invested. We are doing this for four different portfolios to get some comparison. The first method that is used during this comparison is the Mean-variance approach, where we just take the mean of historical data. The second method is the Fama-French model, where we find the βs through Ordinary Least Square method. The third method is also a Fama-French model, but this time we get the βs from quantile regression and Tukey's trimean. The last method is Fama-French as well, and this method is also based on quantile regression but now we use a symmetrical weighting scheme.

The growth of the portfolios

(a)
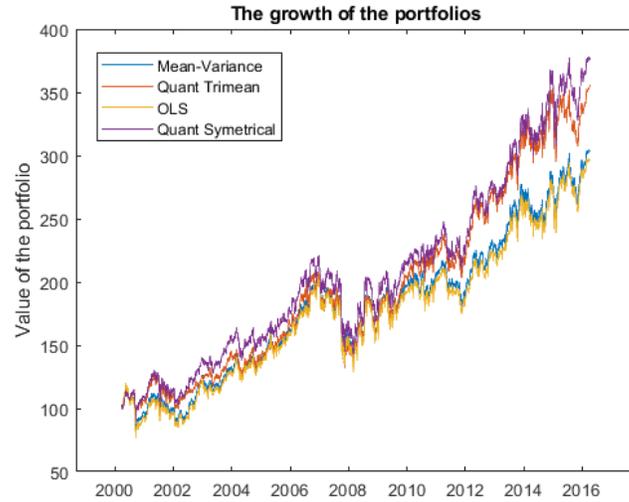
In the figure above we get a picture of how the portfolios are developing over time. We can see that both the quantile based methods are outperforming the Mean-variance and Ordinary Least Square. The symmetrically weighted quantile regression is also outperforming the trimean quantile regression.

We can also have a look at how the yearly performance of each of the portfolios is varying in the table below.

Year	Mean-Variance	OLS	Quant Trimean	Quant Symmetrical
2001	5.84%	2.64%	16.1648 %	18.65%
2002	-9.74%	-9.23%	-12.44%	-10.35 %
2003	20.33%	20.33%	17.62 %	20.14%
2004	11.20%	12.34%	6.12%	10.77 %
2005	11.08%	10.86%	7.53%	5.19%
2006	12.39%	12.71%	16.73%	15.91%
2007	14.21%	14.30%	9.97 %	10.36%
2008	-20.46%	-21.08%	-19.87 %	-20.07%
2009	17.72%	17.71%	17.71%	17.71%
2010	8.80%	8.65%	14.45%	10.82%
2011	4.14%	4.64%	8.60%	8.35%
2012	-1.87%	-2.59%	4.25%	4.31%
2013	7.90%	7.83%	7.42%	5.48%
2014	12.37%	12.95%	15.52 %	15.82%
2015	9.82%	10.19%	9.53%	11.76%
2016	6.13%	6.02%	3.60 %	5.08%

We can see in the table above that Quantile regression with trimean has both the biggest loss and the highest gain in one year. When looking at times of distress, both of the quantile regression methods can be seen to have no direct advantage in 2002. But for both year 2008 and 2012, where in 2012, both the quantile regressions are showing positive return although the other two methods are giving negative return.

Now when we have established that the quantile regression models are outperforming the Mean-variance and the Ordinary Least Square model in creating return. We have to know what kind of risk you are carrying. In the table below have I compiled a table of the standard deviation of these portfolio over this time.

Mean-Variance	Ordinary Least Square	Quantile Trimean	Quantile Symmetrical
10.24%	10.46%	10.55 %	10.57%

In the table above, we can see that there isn't that big of a difference between the risk that the different portfolios are carrying. Which means that the risk-adjusted return are higher when using quantile regression.

8 Portfolio based on CVaR

In the this section will we have a look at how the portfolios vary depending on which risk measure that is used. We will compare variance and the Conditional Value at Risk. The expected return will be calculated by mean of the last year. The results that we get are the following:
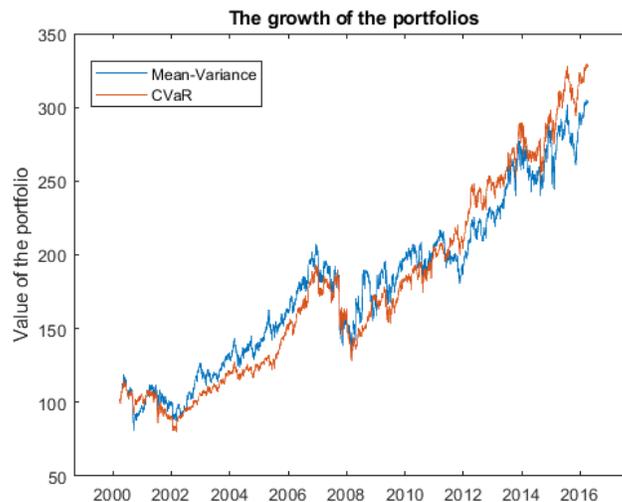
(a)

We can see in the figure above that the Conditional Value at Risk portfolio is outperforming the Mean-Variance portfolio at the end of the period, but this is far from all the time during the whole period. Most of the time is actually the Mean-Variance outperforming the CVaR, it is in the end of the period that CVaR comes ahead of the Mean-Variance. But over the total period the Conditional Value at Risk gets a 25% advantage to the Mean-variance. Then we have to look at the actual risk that we have carried during this time.

Mean-Variance	CVaR
10.24%	11,88%

In the table we have the variance for both the portfolios over this time. We look at the variance for both portfolios since we can not compare Conditional Value at Risk to variance directly. So therefore, we calculate the

variance for the total period of the Conditional Value at Risk portfolio and this gives the table above. We can see that we have carried more risk to receive this additional return and we can have a look at the Sharpe ratio to evaluate how the portfolios have performed.

The Sharpe ratio for the Conditional Value at Risk portfolio was 0.6 and for the Mean-variance portfolio 0.67. This means that we have a better return per unit of risk that we have carried with the Mean-variance portfolio than with the Conditional Value of Risk portfolio.

9   Conclusion

The conclusion that I want to draw from this study is that it can be advantageous in the evaluation of expected return to look at several quantiles, instead of just the central tendency. We can see in the section about how to create expected return, that the returns have a different relation to the market risk premium in the different quantiles. This we can use to our advantage when evaluating the expected return. In times of financial distress, we can see that two out of three times the quantile regression does outperform the Mean-Variance and OLS during this time period. In case of the Conditional Value at Risk portfolio, even if we don't get a higher Sharpe ratio, we do get a higher return over the whole period, with the same expected return. But this is very ambiguous since the CVaR is under performing the Mean-Variance portfolio during this time period. So we can not draw any conclusion of which of these risk measures that are the better one in portfolio optimization, this would require more empirical tests.

References

[1] Acerbi, Carlo, Tasche Dirk, *On the coherence of expected shortfall*, Journal of Banking & Finance, Vol. 26, 1487-1503, 2002.

[2] Allen, D. E, powell, R. J and Singh A. K, *Quantile Regression as a Tool for Portfolio Investment Decisions During Times of Financial Distress*, Annuals of Financial Economics, vol 6, No 1, 2011.

[3] Kaplan, Paul D, *Frontier of Modern Asset Allocation*, Morningstar, 2011.

[4] Kisiala, Jakob, *Conditional Value at Risk: Theory and Applications*, The university of Edinburgh, 2015.

[5] Klimek, Maciej, *Lecture note: Basic information about quantile regression and VaR* , Uppsala University, March 2018.

[6] Krokhmal, P, Palmquist, J and Uryasev, S, *Portfolio optimization with conditional value-at-risk objective and constraints*, 2001

[7] Luenberger, David G, *Investment Science*, Oxford University Press, 1998.

[8] Savage, Sam, *The Flaw of Averages*, John Wiley & Sons, 2009.

[9] Tukey, John Wilder, *Exploratory Data Analysis*, Addison-Wesley, 1977.