



Norwegian University of  
Science and Technology

# Risk Parity Stock Optimization Using Principal Component Quantile Simulation

**Bård Ø. Aanonsen**  
**Anders Loe Fagerholt**

Industrial Economics and Technology Management

Submission date: May 2016

Supervisor: Sjur Westgaard, IØT

Norwegian University of Science and Technology  
Department of Industrial Economics and Technology Management



## Problem Description

Estimating future asset returns is a difficult problem making input sensitive portfolio optimization models, like the mean-variance model, perform poorly when comparing risk adjusted returns. Many investors still use the traditional 60/40 rule as it offers a degree of diversification and often beats other portfolio optimization heuristics in the financial markets.

The recent popularity of the risk parity portfolio heuristics has given a new hope for technical optimization models as this heuristic only focuses on risk contributions, not expected returns, making it more robust. By equalizing risk contributions, true diversification is offered achieving higher risk adjusted return. So far volatility has been the preferred risk measure in research, but it has its limitations.

This thesis will develop a new framework for simulating multivariate return distributions to be used in a risk parity portfolio optimization model. By introducing a modern risk measure in risk parity optimization, better intelligence is offered to investors and less risky portfolios may be obtained.



## Preface

This thesis is written the Spring of 2016 as the concluding part of our five year master's degree programme in industrial economics and technology management at The Norwegian University of Science and Technology. We have experience in financial engineering and industrial optimization from The Department of Industrial Economics and Technology Management (IØT) making this a unique opportunity for us to gain more experience and synthesize our knowledge from both fields.

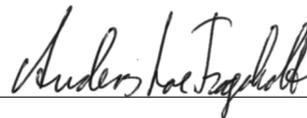
We would like to thank our supervisor Professor Sjur Westgaard from IØT for ideas, insight and inspiration along the way. His supervision and thorough guidance has been invaluable throughout the semester.

Trondheim, June 2016



---

Bård Aanonsen



---

Anders Loe Fagerholt



## Sammendrag

Dagens porteføljeoptimeringsmodeller er ofte for sensitive til stokastiske inputparametre og bruken av utdaterte risikomål. Som følge av dette lider mange av dem av dårlig risikojustert avkastning. Denne avhandlingen presenterer en løsning som unngår disse problemene ved å introdusere et nytt rammeverk for å simulere multivariate avkastninger for korrelerte aktiva. I tillegg introduseres en modell for å konstruere en risikoparitetsportefølje med *forventet haletap* (CVaR) som risikomål. Dette gjør at modellen kan utnytte spesifikk haleinformasjon funnet i marginaldistribusjonene.

*Kvantilregresjon* og *prinsipialkomponentanalyse* brukes til å formulere en faktormodel som modellerer hele distribusjonen og samtidig opprettholder avhengigheter mellom aktiva. En ny måte å simulere framtidige prinsipialkomponenter på gjør simuleringsalgoritmen rask og effektiv.

De resulterende marginaldistribusjonene ivaretar aktivaspesifikke risikokarakteristikker og haleoppførsel. Dette gjenspeiles tydelig i riskoparitetsporteføljen, noe som viser at CVaR som risikomål tilbyr rikere informasjon til investorer.



## Abstract

Today's portfolio optimization models are often too sensitive to stochastic input parameters and the use of outdated risk measures, resulting in poor risk adjusted return. This thesis presents a solution to avoid these difficulties by introducing a new simulation framework to obtain multivariate return distributions for correlated assets. Next a model to obtain a risk parity portfolio using *Conditional Value at Risk* (CVaR) is offered making the model able to capture asset specific risk characteristics present in the tails of the marginal distributions.

*Quantile regression* and *principal component analysis* (PCA) are combined to form a factor model able to capture the entire return distribution and maintain dependencies between correlated assets. A new method to simulate future principal components is presented making the simulation algorithm quick and effective.

The resulting marginal distributions show asset specific risk characteristics and tail behaviour. This is in turn reflected in the risk parity portfolio weights, confirming CVaR as a risk measure offering better intelligence to investors.



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Methodology</b>	<b>5</b>
2.1	Factor Model . . . . .	5
2.1.1	The Arbitrage Pricing Theory . . . . .	6
2.1.2	Quantile Regression . . . . .	7
2.1.3	Principal Component Analysis . . . . .	8
2.1.4	PCA Factor Model . . . . .	11
2.2	Simulation of Return Distributions . . . . .	13
2.2.1	The Simulation Algorithm . . . . .	13
2.2.2	Simulation of Principal Components . . . . .	15
2.2.3	Construction of Future Return Distributions . . . . .	24
2.3	Risk Parity Portfolio Optimization . . . . .	27
2.3.1	Mathematical Representation . . . . .	27
2.3.2	The Optimization Algorithm . . . . .	28
2.3.3	Risk Measures in Risk Parity Optimization . . . . .	30
<b>3</b>	<b>Empirical Results</b>	<b>35</b>
3.1	Data Description . . . . .	35
3.2	Empirical Model . . . . .	37
3.3	Quantile Regression Results . . . . .	40
3.4	Simulated Cumulative Return Distributions . . . . .	43
3.5	Simulated Return Distributions . . . . .	45
3.6	Simulation Performance . . . . .	48
3.7	Portfolio Optimization . . . . .	52

3.7.1	Optimization Results . . . . .	53
<b>4</b>	<b>Conclusion</b>	<b>57</b>
<b>A</b>	<b>Computational Information</b>	<b>63</b>
<b>B</b>	<b>Data set</b>	<b>64</b>
<b>C</b>	<b>Descriptive Statistics of Historical Data</b>	<b>67</b>
<b>D</b>	<b>Sector Comparison</b>	<b>72</b>
D.1	Variance Explained by PC . . . . .	72
D.2	Correlation Matrices . . . . .	73
<b>E</b>	<b>Explanation of Numerical Instability</b>	<b>74</b>
<b>F</b>	<b>January 2016 CI Analysis</b>	<b>75</b>
<b>G</b>	<b>Simulation Plots</b>	<b>77</b>
<b>H</b>	<b>Risk Parity Portfolio Weights</b>	<b>79</b>

## List of Figures

1	Quantile regression lines and OLS . . . . .	8
2	Orthogonal transformation in PCA . . . . .	9
3	Scatter plot of calculated cumulative return distribution for a stock given set of principal components . . . . .	12
4	Flow chart of the simulation algorithm . . . . .	15
5	Simulated return from a cumulative return distribution . . . .	16
6	4 simulations of correlated asset movements comparing simu- lated principal components . . . . .	23
7	AAPL's distribution of monthly returns from 1995 to 2015 . .	25
8	Development of return distributions as the number of simula- tions increases . . . . .	26
9	Risk measures discussed . . . . .	32
10	Comparison of correlation matrices between data sets . . . . .	37
11	Effect of PCA, comparison between sector 10, 25, 40 and the full data set . . . . .	39
12	The cumulative distribution with too many explanatory vari- ables . . . . .	41
13	$\alpha$ and $\beta$ estimates from quantile regression . . . . .	42
14	Simulated cumulative stock return distributions for APA, XOM, ABT and LLY . . . . .	44
15	Simulated stock return distributions for January 2016 . . . . .	46
16	Monthly development of the S&P500 index from 1995 to Q2 2016. . . . .	47

17	Forecast plot for January 2016 using probability distributions from figure 15 . . . . .	50
18	Realized and mean returns with 95% confidence intervals . . .	51
19	Example portfolio of two stocks with different characteristics. .	52
20	Simulated distributions and risk parity weights for sector 10 .	55
21	Cumulative variance and variance explained by principal com- ponent for all sectors and the full data set, respectively . . . .	72
22	Correlation matrices for all sectors and for the full data set . .	73
23	Realized and mean returns . . . . .	76
24	Return distributions for all stocks with 100.000 simulations . .	78
24	Risk parity weight allocations for all sectors using three dif- ferent risk measures . . . . .	80

## List of Tables

1	Effect of PCA with at least 95% variance explained . . . . .	38
2	Descriptive statistics of probability distributions from figure 15	46
3	Confidence intervals of probability distributions from figure 15	49
4	Relative amount in CI Intervals . . . . .	51
5	Target function values for the different risk measures used in the portfolio optimization . . . . .	54
6	Full data set . . . . .	67
7	Descriptive statistics of Historical Data . . . . .	71



# 1 Introduction

Being able to accurately forecast the price movements of stocks and other assets would give a superior advantage in today's financial markets. This problem is hard to solve and with markets being in constant change, a promising solution today might not be as promising tomorrow. Merton concluded in 1980 that estimating the market's return is difficult to do with much accuracy, as the expected return is too sensitive to changing factors [1]. This does not only apply to the market's return, but applies to many stochastic variables. As a consequence of this, portfolio optimization is still a problem whose solutions may look good on paper, but perform poorly when implemented in practice.

Markowitz [2] presented a solution to the portfolio optimization problem using a solid mathematical foundation with his *mean-variance* (MPT) framework in 1952. Since then many researchers have extended this theory which has led to the introduction of several heuristics such as Arnott, Hsu & Moore's *fundamental weights* (FW) [3], CAPM's *market value weights* (MVW) and the MPT special case, *minimum variance* (MV) [4]. With all these alternatives at hand, many practitioners still use the traditional 60/40 rule because the heuristics have not proven to be superior to the 60/40 portfolio in financial markets. One of the reasons for this is that the models base their output on stochastic parameters, which are difficult to predict, such as an asset's future return. That being said, trying to predict an asset's future movement may not be a waste of time. Even though an asset's expected return is difficult to estimate, the shape of the return distribution is easier to estimate and can support the decision making in several applications

including portfolio optimization.

Risk parity is a new portfolio heuristic with great potential as it does not rely on the estimates of future asset prices. Its only objective is to equalize the total risk contribution from every asset, thus creating the highest degree of diversification achievable. The general idea was developed in the 1990s, but after the financial crisis in 2008, it raised researchers' attention. Maillard, Roncalli & Teiletche [6] introduce the properties of the heuristic and how to create a risk parity portfolio, and Denis Chaves, Jason Hsu, Feifei Li & Omid Shakernia [7] claim the risk parity portfolio indeed offers true diversification while showing promising risk adjusted return compared to other heuristics, including the 60/40 portfolio.

Another powerful characteristic of the risk parity portfolio is that the investor may choose any risk measure to compare the risk contributions, making future asset return distributions highly relevant. So far, volatility has been the risk measure in focus among researchers, but for the purpose of portfolio optimization, it is an outdated risk measure. *Conditional Value at Risk* (CVaR) also known as *expected tail loss* (ETL) or *expected shortfall* (ES), captures the tail information in the return distributions providing better intelligence to the investors. Therefore it is a more expedient risk measure in a state of the art portfolio optimization model.

This thesis presents a new framework to simulate multivariate return distributions of correlated assets while keeping important asset specific tail information. It then shows how to use these distributions in creating a risk parity portfolio, with a risk measure that captures this information. Chapter 2 establishes this framework and its building blocks. In chapter 3 the main

objective is to show how the framework is applied, what type of information one can extract from the results and how one uses the information to obtain a risk parity stock portfolio. This way an understanding of how the framework operates, what information to expect and how to use this information in a portfolio optimization model is offered.

*Arbitrage pricing theory, principal component analysis (PCA) and quantile regression* are combined to create a multi-factor model able to describe future scenarios. This model along with a comprehensive simulation algorithm based on Alexander & Ruppert's algorithm [5], results in multivariate asset return distributions which are used in the risk parity portfolio optimization model.

This thesis makes a valid contribution to financial research as it presents a new algorithm to simulate multivariate return distributions and obtain a risk parity portfolio using a suitable risk measure. Focusing entirely on risk, these methods are combined in an original framework, shifting the attention from estimating future mean returns to simulating future risk characteristics. This way one avoids some of the difficulties present in existing portfolio optimization models.

Chapter 4 concludes the thesis, discusses future research and possible extensions of this framework.



## 2 Methodology

This chapter introduces the framework used to simulate multivariate return distributions and the general process to obtain a risk parity portfolio using the simulated stock returns. The chapter consists of three sections where section 2.1 focuses on the mathematical foundation of the framework, section 2.2 presents the mechanics behind the simulation algorithm and lastly section 2.3 covers the implementation of the simulation results in risk parity optimization. Throughout chapter 2, a brief overview of the mathematical building blocks is given to enlighten the reasoning behind the framework before the models are discussed.

### 2.1 Factor Model

The most popular financial factor model is the *capital asset pricing model* (CAPM) first introduced by Sharpe [8] and Lintner [9] in the 1960s. This model identifies the undiversifiable risk of an asset and illustrates the linear relationship between the asset's expected return given the excess return of the market. This relationship can be modeled by the following factor model.

$$r_i = r_f + \beta_i R_M + \epsilon_i, E(\epsilon_i) = 0, \quad (1)$$

The constant term  $r_f$  is the risk free interest rate and  $\beta_i$  measures the asset's sensitivity to the market's excess return,  $R_M$ , which is the return over the risk free rate. A positive beta will indicate a positive correlation between  $i$  and the market, and vice versa. In addition, the beta's value will indicate high or low volatility given changes in the market's excess return. A solution

to this model can be found by minimizing  $\epsilon_i^2$  which is called the squared residual or squared error. The most used method to solve this optimization problem is *ordinary least squares* (OLS).

### 2.1.1 The Arbitrage Pricing Theory

In the financial markets factor models play an important part in modeling a relationship between a risky asset and different external risk factors. Risk factors can be the return of a market index, an interest rate or other macro economic variables. These models are in a subcategory of the *arbitrage pricing theory* (APT) developed by Ross [10]. How an asset behaves given the behaviour of a risk factor is a question that can be explained by factor models. These models often contain multiple risk factors and are then called *multi-factor* models. The general APT model takes the vector form,

$$r_i = \alpha_i + \beta_{1,i} F_1 + \dots + \beta_{n,i} F_n + \epsilon_i \quad (2)$$

here the return of a risky asset  $i$ , is described by a linear combination of  $n$  risk factors including a constant term and an error term. The constant term is constant because it is riskless and not influenced by variable risk factors. The betas are called *factor sensitivities* and is a measure of how sensitive the risky asset is to the known factor. If the beta is far from zero,  $i$  will be heavily influenced by the factor, thus even a small change in the risk factor could have a great impact on the risky asset. The alpha and the betas can be estimated using OLS. These types of models are very useful because it is often easier to measure and predict the future movements of risk factors than it is to predict an asset's return. If one can show a relationship between

the movements of an asset's return and risk factors, and at the same time be able to predict the movement of the same risk factors, a factor model would be an efficient tool to predict the risky asset's movement.

### 2.1.2 Quantile Regression

Solving equation 2 using OLS will output the expected  $r_{i,t}$  at a given point in time conditional on a set of factors, that is

$$E(r_{i,t}|F_{1,t}, \dots, F_{n,t}) = \alpha_i + \beta_{i,1}F_{1,t} + \dots + \beta_{i,n}F_{n,t} \quad (3)$$

Because  $r_{i,t}$  is a stochastic variable, it has a probability distribution. As this method finds the conditional mean it does not capture any information regarding this distribution. Quantile regression introduced by Koenker & Bassett [11], is similar to OLS, but the regression line is conditional on a specified quantile. Consider two time series X and Y of T observations, with a set of  $(x_t, y_t)$  pairs OLS can be applied on the set and produce  $E(y_t|x_t)$ . With quantile regression a similar expression can be produced only here the regression line will divide the set of  $(x_t, y_t)$  pairs at the q-quantile. Figure 1 illustrates a scatter plot with  $(x_t, y_t)$  pairs and the same plot with three regression lines, OLS, the 0.1-quantile and the 0.9-quantile regression line.

The true power of quantile regression becomes available when applying a full set of quantiles in the range  $\langle 0, 1 \rangle$ . Defining the q-quantile regression line for the return of stock  $i$  as,

$$Q_{r_i}(q|F_1, \dots, F_n) = \alpha_i^q + \beta_{1,i}^q F_1 + \dots + \beta_{n,i}^q F_n + \epsilon_i^q \quad (4)$$

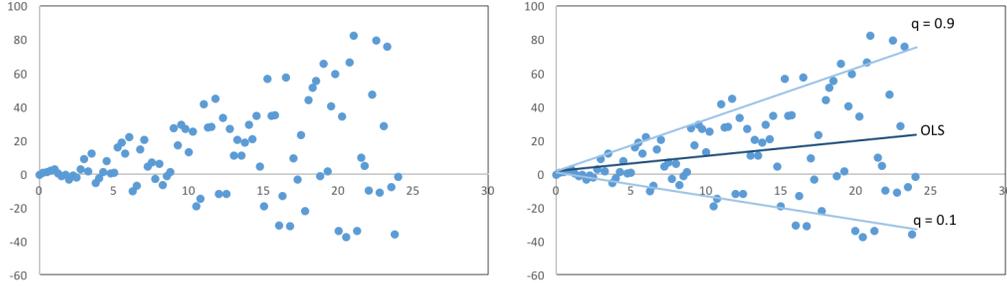


Figure 1: Quantile regression lines and OLS

multiple regression lines for a set of quantiles can be used to create a cumulative probability function for  $r_{i,t}$ . In equation 4 the variables  $r_i$ ,  $F_j$  and  $\epsilon_i$  are time series represented as column vectors with T observations.

### 2.1.3 Principal Component Analysis

Principal component analysis is an orthogonalization technique to transform a set of correlated variables into a set of uncorrelated variables. Figure 2 depicts this orthogonal linear transformation in two dimensions. The uncorrelated variables are called principal components and are indeed orthogonal because they are computed from the normalized eigenvectors from the covariance matrix.

These correlated variables are in this case time series of asset returns. PCA uses the properties of the time series' covariance matrix and outputs a set of principal components which can easily reproduce the original data set. Taking a data set represented by a  $T \times n$  matrix  $\mathbf{X}$  of  $n$  assets with  $T$  observations through time with a covariance matrix  $\mathbf{C}$ , the principal components of  $\mathbf{X}$  can be found by multiplying the data set with an  $n \times n$  matrix  $\mathbf{w}$ ,

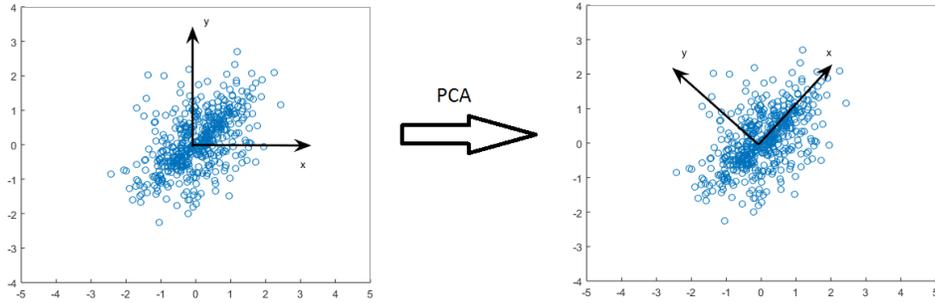


Figure 2: Orthogonal transformation in PCA

$$\mathbf{P} = \mathbf{X}\mathbf{w} \quad (5)$$

where the columns of  $\mathbf{w}$  are the normalized eigen vectors of  $\mathbf{C}$  and the columns of  $\mathbf{P}$  are the principal components. The columns of  $\mathbf{w}$  are sorted so that the first column vector is the eigen vector with the largest corresponding eigenvalue, that is the vector which explains most of the variance in  $\mathbf{X}$ . The second column vector is the eigen vector explaining the second most variance in  $\mathbf{X}$  and so forth. Put differently, for all  $n$   $(\lambda_i, \mathbf{E}_i)$  pairs of  $\mathbf{C}$ , where  $\lambda_i$  is the eigen value and  $\mathbf{E}_i$  is the corresponding normalized eigen vector,  $\mathbf{w}$  is given by equation 6

$$\mathbf{w}_{n \times n} = \left( \mathbf{E}_1 \dots \mathbf{E}_n \right) \quad (6)$$

where,

$$\lambda_1 \geq \dots \geq \lambda_n \quad (7)$$

holds.

When equation 6 and 7 holds, the variation in the data set explained by the  $i$ th principal component is given by

$$\frac{\lambda_i}{\lambda_1 + \dots + \lambda_i + \dots + \lambda_n} \quad (8)$$

This equation can be used to calculate how many principal components are needed to explain the variation in the data set up to a given threshold. All  $n$  principal components explain 100% of the variation in  $\mathbf{X}$ , but  $m$  principal components where  $m < n$  might be able to explain for example 95% of the variance. If 95% is a good enough margin, all columns from  $m+1$  and up to  $n$  can be removed from  $\mathbf{w}$ . For a highly correlated set of variables,  $m \ll n$  and the benefit of using PCA is higher. The original data set  $\mathbf{X}$  can be reconstructed using

$$\mathbf{X} = \mathbf{P}\mathbf{w}^T \quad (9)$$

Because  $\mathbf{w}$  is orthogonal,  $\mathbf{w}^{-1} = \mathbf{w}^T$ . With a 95% variance margin, only  $m$  columns are needed and  $n - m$  columns can be removed from  $\mathbf{w}$ . If  $\mathbf{w}^*$  is the matrix with the  $m$  leftmost columns of  $\mathbf{w}$ ,  $\mathbf{P}^*$  can be found using equation 5 with  $\mathbf{w}^*$  instead of  $\mathbf{w}$ .  $\mathbf{P}^*$  is then the matrix with the  $m$  first principal components.

$$\mathbf{X} \approx \mathbf{P}^*\mathbf{w}^{*T} \quad (10)$$

With  $\mathbf{P}^*$  and  $\mathbf{w}^*$ ,  $\mathbf{X}$  can be approximated keeping 95% of the original variance with a reduction of  $n - m$  dimensions. This is why PCA is such a powerful tool to reduce complexity of large correlated data sets. When the framework is used to simulate correlated asset returns, the effect of using

PCA should increase as higher correlation in the data set leads to a higher effect of PCA.

#### 2.1.4 PCA Factor Model

With the mathematical foundations explained, the original model for representing asset returns can be introduced. This model is a multi-factor model which makes use of quantile regression and principal component analysis to represent cumulative return distributions. The input to the factor model is a data set  $\mathbf{X}$ , with T observations of n assets. PCA outputs a reduced set of principal components large enough to explain a set margin of the variance in the original data set. The PCA factor model is then given by equation 11.

$$Q_{r_i}(q|PC_1, \dots, PC_m) = \alpha_i^q + \beta_{1,i}^q PC_1 + \dots + \beta_{m,i}^q PC_m + \epsilon_i^q \quad (11)$$

and is run for all n stocks for all q quantiles. The quantiles are,

$$Q = \{0.01, 0.1, 0.2, \dots, 0.9, 0.99\} \quad (12)$$

The 0.01-quantile and the 0.99-quantile are included to capture tail behaviour in the distributions. For a given set of principal components,  $\mathbf{PCs}$ , one can calculate  $Q_{r_i}(q|\mathbf{PCs})$  for all quantiles and get a scatter plot. An example is illustrated in figure 3. Notice how the first and last quantile explain important tail behaviour and make the distribution look like an S-curve. These data points can later be interpolated to create a full cumulative return distribution used for asset return simulation.

As mentioned in the introduction, the objective behind this simulation framework is to capture dependencies between assets and simulate their

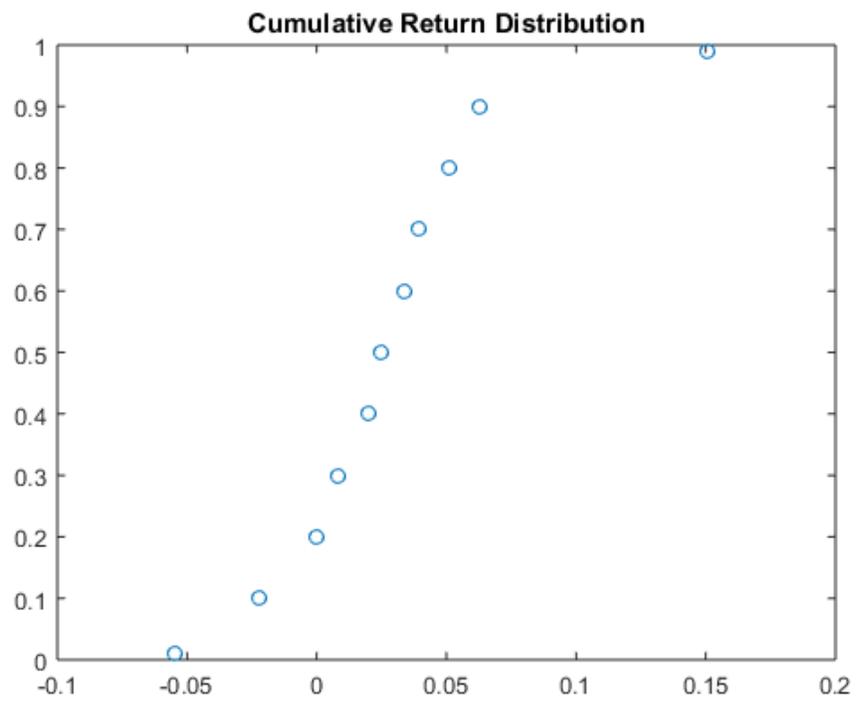


Figure 3: Scatter plot of calculated cumulative return distribution for a stock given set of principal components

marginal distributions. With **PCs** as explanatory variables, these dependencies are easier to model because of the variance properties of the principal components explained in section 2.1.3. Quantile regression makes it possible model the entire distribution instead of the conditional mean. By combining these methods, one gets a factor model able to fulfill the requirements described in chapter 1.

## **2.2 Simulation of Return Distributions**

The simulation algorithm combines principal component analysis, quantile regression, factor model theory and numerical interpolation to produce a probability distribution for the return of an asset at the end of a specified period. The core of the simulation framework is an extension of Alexander & Ruppert [5], however in this simulation a more extensive factor model is developed and used to increase efficiency and better model the dependencies. Section 2.2.1 introduces the simulation algorithm with a simple layout in order to get a good overall understanding of how the algorithm works. In later sections more technical details will be presented and explained.

### **2.2.1 The Simulation Algorithm**

Figure 4 shows a flow chart of the simulation algorithm. The algorithm is made up of four main procedures, where one of these is explained in a sub chart. The first procedure is the principal component analysis on the historical data. Here  $\mathbf{w}$  from equation 5 is calculated and the unnecessary columns are removed before it is multiplied with the matrix containing the historical data. This procedure results in a reduced set of principal components which

is able to explain as much variance as required. The principal components are then used as explanatory variables in the factor model. The next step is to estimate an  $\alpha$  and multiple  $\beta$ s for each quantile, for each asset in the data set. These constants are the ones in equation 11, and quantile regression is used to find these estimates. The first phase of the algorithm is now complete. It involves everything that have to do with looking back in time and processing historical data.

Phase two is about predicting the future and starts by simulating a set of principal components. This procedure is the first part of a bigger loop that must be done many times to improve the accuracy of the return estimates. With simple pre-allocation of memory and input, this loop can easily run in parallel on a computer or on a cluster network to significantly reduce the time it takes to simulate the future return distributions <sup>1</sup>. With a simulated set of future principal components and the  $\alpha$ s and  $\beta$ s from the quantile regressions, equation 11 is used to calculate a data point for each quantile. The set of data points produces a scatter plot of the cumulative return distribution for the asset. This procedure is done for every asset. The mentioned scatter plot is already illustrated in figure 3. In the next step these data points are interpolated to create a continuous function describing the cumulative return distribution.

With all this done, the last step consists of generating a uniform random number between 0 and 1 and drawing a return from the cumulative return distribution. This drawn return, will represent a simulated future asset return. By repeating this process thousands of times, these drawn returns will

---

<sup>1</sup>See appendix A for computational information.

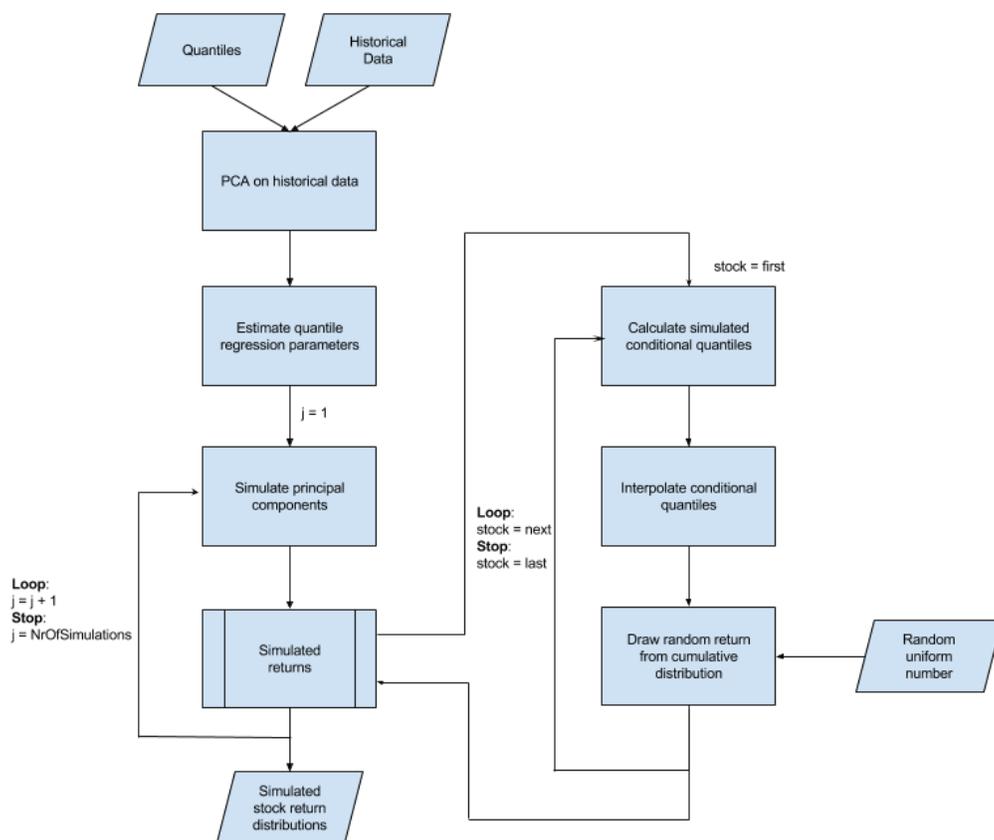


Figure 4: Flow chart of the simulation algorithm

form a probability distribution for the future return of the asset. The last two steps are illustrated in figure 5. This concludes phase two and the entire simulation algorithm <sup>2</sup>.

### 2.2.2 Simulation of Principal Components

The simulation of principal components is the most important part of the entire stock simulation algorithm as the simulated future returns are depen-

<sup>2</sup>See algorithm 1 for a consolidated description in pseudo code.

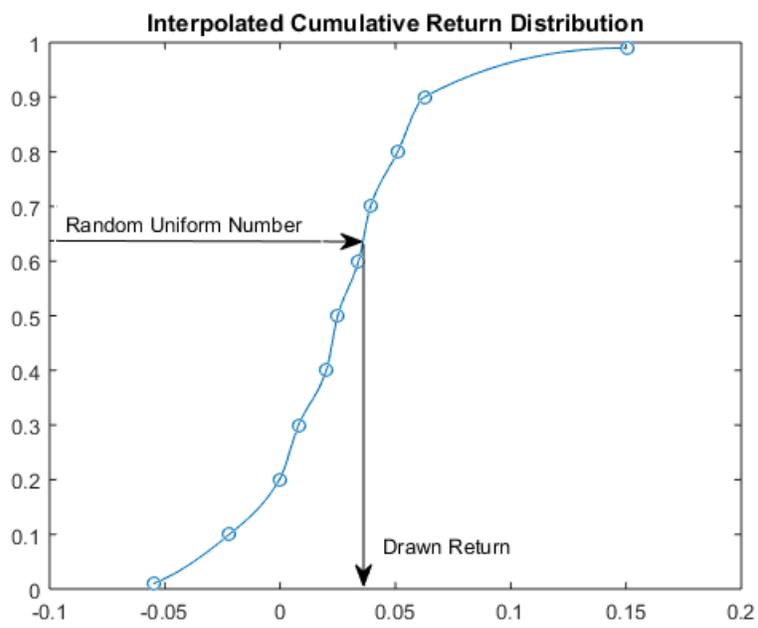


Figure 5: Simulated return from a cumulative return distribution

**input** : A set of quantiles  $Q$   
**input** : A  $T \times n$  matrix  $\mathbf{X}$ , of historical data for  $n$  assets  
**input** :  $\tilde{S}$ , number of simulations  
**output**: An  $\tilde{S} \times n$  matrix  $\mathbf{D}$ , of simulated returns

*Perform PCA and obtain the reduced set of principal components, PCs, of  $\mathbf{X}$  where the variance margin hold;*

```

for  $i \leftarrow 1$  to  $n$  do
    foreach  $q$  in  $Q$  do
        Estimate the parameters  $\alpha_i^q, \beta_{1,i}^q, \dots, \beta_{m,i}^q$  from equation 11;
    end
end
for  $j \leftarrow 1$  to  $\tilde{S}$  do
    Simulate a set of future principal components, simPCs;
    for  $i \leftarrow 1$  to  $n$  do
        Calculate the conditional quantiles from equation 11 using simPCs;
        Interpolate the conditional quantiles to obtain a cumulative return distribution;
        Simulate a return by drawing from the distribution and insert into  $\mathbf{D}(j, i)$ ;
    end
end

```

**Algorithm 1:** The Simulation Algorithm

dent upon the simulated future principal components through the PCA factor model shown in equation 11. The original idea behind these future principal components comes from equation 5. When a matrix of principal components can be found by multiplying  $\mathbf{w}$  with the matrix of historical returns, the same logic could be applied into the future. Meaning a simulated matrix of principal components  $\tilde{\mathbf{P}}$  can be found by multiplying  $\mathbf{w}$  by a simulated set of future returns  $\tilde{\mathbf{X}}$ , that is,

$$\tilde{\mathbf{P}} = \tilde{\mathbf{X}}\mathbf{w} \quad (13)$$

By simulating asset movements one can create  $\tilde{\mathbf{X}}$ . One of the ways to do this is by modeling movements as stochastic processes, treating the returns over the time span as a system of random variables.

$$\tilde{\mathbf{P}} = \underbrace{\begin{pmatrix} \textit{stochastic} \\ \textit{processes} \end{pmatrix}}_{\tilde{\mathbf{X}}} \mathbf{w} \quad (14)$$

This idea makes one assumption that  $\mathbf{w}$  is constant in the simulated future. As the matrix  $\mathbf{w}$  is calculated based on the historical data, this is a reasonable assumption to make. In chapter 3 the plain stochastic process for simulating stock movements is used, that is *geometric Brownian motion* (GBM). However, one could use any relevant stochastic process to simulate asset movements, like a jump-diffusion process or other processes specialized in modeling the chosen asset's behaviour.

## Geometric Brownian Motion For Correlated Stocks

For a single stock the proportional change of the price at time  $t$  can be written,

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t \quad (15)$$

This process is called a geometric Brownian motion because the change in price is relative to  $S_t$  compared to an absolute change which is an *arithmetic Brownian motion*.  $\mu$  is the stocks drift,  $\sigma$  is the volatility over the period and  $W_t$  is a Brownian motion. The solution to this stochastic differential equation is, [12]

$$S_t = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t} \quad (16)$$

To simplify the math for computing purposes the numerical approximation of this stochastic differential equation is of interest to the simulation algorithm. For a discrete step in time equation 16 becomes,

$$S_{t+\Delta t} = S_t e^{(\mu - \frac{\sigma^2}{2})\Delta t + \sigma W_{t+\Delta t}} \quad (17)$$

Here  $\mu$  and  $\sigma$  are adjusted for the period  $\Delta t$ . Using the *Euler-Maruyama method*, equation 17 becomes,

$$S_{t+\Delta t} = S_t + \mu S_t \Delta t + \sigma S_t \Delta W_t \quad (18)$$

where  $\Delta W_t = W_{t+\Delta t} - W_t$  and  $\Delta W_t \sim N(0, \Delta t)$ . Notice how the price term is included in all right hand parts of equation 18, because the algorithm is focused on returns, this equation can be simplified to,

$$\tilde{r}_{t+\Delta t} = \frac{S_{t+\Delta t} - S_t}{S_t} = \mu + \sigma \Delta W_t \quad (19)$$

assuming  $\Delta t = 1$ .

Stocks in the same markets and sectors tend to move together and are influenced by each other. For this reason, GBM of several correlated assets should be simulated as a system. Equation 19 for a system of  $n$  correlated assets becomes,

$$\tilde{r}_{t+\Delta t} = \underbrace{\mu}_{1 \times n} + \underbrace{\Delta W_t}_{1 \times n} \underbrace{\mathbf{L}^T}_{n \times n} \quad (20)$$

where  $\mathbf{L}$  is the lower *Cholesky decomposition* of the covariance matrix  $\mathbf{C}$  and  $\Delta W_t \sim MVN(0, I)$ .

### Principal Component Simulation with GBM

With an understanding of geometric Brownian motion for correlated stocks, the framework's simulation of principal components can be explained. If the algorithm is run on other assets than stocks or an other stochastic process than GBM, the following procedure must be altered accordingly.

With GBM  $\tilde{\mathbf{P}}$  is created the following way,

$$\tilde{\mathbf{P}} = \underbrace{\begin{pmatrix} \mu + \Delta W_t \mathbf{L}^T \\ \mu + \Delta W_{t+1} \mathbf{L}^T \\ \vdots \\ \mu + \Delta W_{t+\tilde{T}} \mathbf{L}^T \end{pmatrix}}_{\tilde{\mathbf{x}}} \mathbf{w} \quad (21)$$

Here the principal components are simulated  $\tilde{T}$  steps into the future where

$\Delta t = 1$ . With this method, an assumption is made that  $\mu$ , which are calculated based on historical data, holds in the future. By exploiting the benefits of PCA and only using the  $m$  principal components needed to fit the requirement of variance explained, one can experience a reduction in computational effort in the calculation of future principal components by multiplying  $\tilde{\mathbf{X}}$  by  $\mathbf{W}^*$  instead of  $\mathbf{W}$  getting,

$$\tilde{\mathbf{P}}^* = \tilde{\mathbf{X}}\mathbf{w}^* \quad (22)$$

With this method  $\tilde{\mathbf{X}}$  still needs to be fully computed using GBM or another stochastic process before multiplying it with  $\mathbf{w}^*$ . Because  $\tilde{\mathbf{X}}$  has to be simulated in every simulation, it is a time consuming step. In the end, the goal is to end up with a matrix of simulated principal components  $\tilde{\mathbf{P}}^*$ . For an arbitrary row in the matrix,

$$\tilde{\mathbf{p}}_t^* = \tilde{\mathbf{x}}_t \mathbf{w}^* \quad (23)$$

$1 \times m$              $1 \times n$   $n \times m$

where,

$$E[\tilde{\mathbf{p}}_t^*] = E[\tilde{\mathbf{X}}_t \mathbf{w}^*] = \mu \mathbf{w}^* \quad (24)$$

and

$$Var[\tilde{\mathbf{p}}_t^*] = \mathbf{w}^{*T} \mathbf{L} \mathbf{L}^T \mathbf{w}^* = \mathbf{w}^{*T} \mathbf{C} \mathbf{w}^* = \tilde{\mathbf{C}}_{m \times m} \quad (25)$$

To make the algorithm more efficient in simulating future principal components, it splits up  $\tilde{\mathbf{X}}$  and performs the GBM simulation and PCA multiplication from equation 23 in one calculation using the following equation,

$$\underset{1 \times m}{\tilde{\mathbf{P}}_t^*} = \underset{1 \times n}{\mu} \underset{n \times m}{\mathbf{w}^*} + \underset{1 \times m}{\Delta W} \underset{m \times m}{\tilde{\mathbf{L}}^T} \quad (26)$$

where  $\tilde{\mathbf{L}}$  is the lower Cholesky decomposition of  $\tilde{\mathbf{C}}$ , which is static. Notice how the random vector  $\Delta W$  can be reduced from a length  $n$  to  $m$  using this method. To show that this works,  $\tilde{\mathbf{p}}_t^*$  calculated by equation 26 must have the same expected value and variance as in equation 24 and 25.

$$E[\tilde{\mathbf{p}}_t^*] = E[\mu \mathbf{w}^*] + E[\Delta W_t \mathbf{L}^T \mathbf{w}^*] = \mu \mathbf{w}^* + 0 \quad (27)$$

and variance,

$$Cov[\tilde{\mathbf{P}}_t^*] = Cov[\mu \mathbf{w}^*] + Cov[\Delta W \underbrace{\mathbf{L}^T \mathbf{w}^*}_A] = \mathbf{A}^T Cov[\Delta W] \mathbf{A} \underset{(eq29)}{=} \tilde{\mathbf{C}} \quad (28)$$

$$\mathbf{A}^T Cov[\Delta W] \mathbf{A} = \mathbf{w}^{*T} \mathbf{L} \mathbf{L}^T \mathbf{w}^* = \mathbf{w}^{*T} \mathbf{C} \mathbf{w}^* = \tilde{\mathbf{C}} \quad (29)$$

Using this process, a significant amount of computational time is saved making the simulation algorithm faster.

Figure 6 shows a comparison between simulating principal components. The three columns use equation 13, 22 and 26 respectively, and each row shows a simulation run. Here the future movement of six arbitrary assets are shown using returns calculated with equation 9 and 10 from simulated principal components. Notice how the trend is clearly similar between the plots, this is caused by the equality between equation 24 and 27. The covariance matrix of each plot in a simulation should also be equal as a consequence of the variance equality.

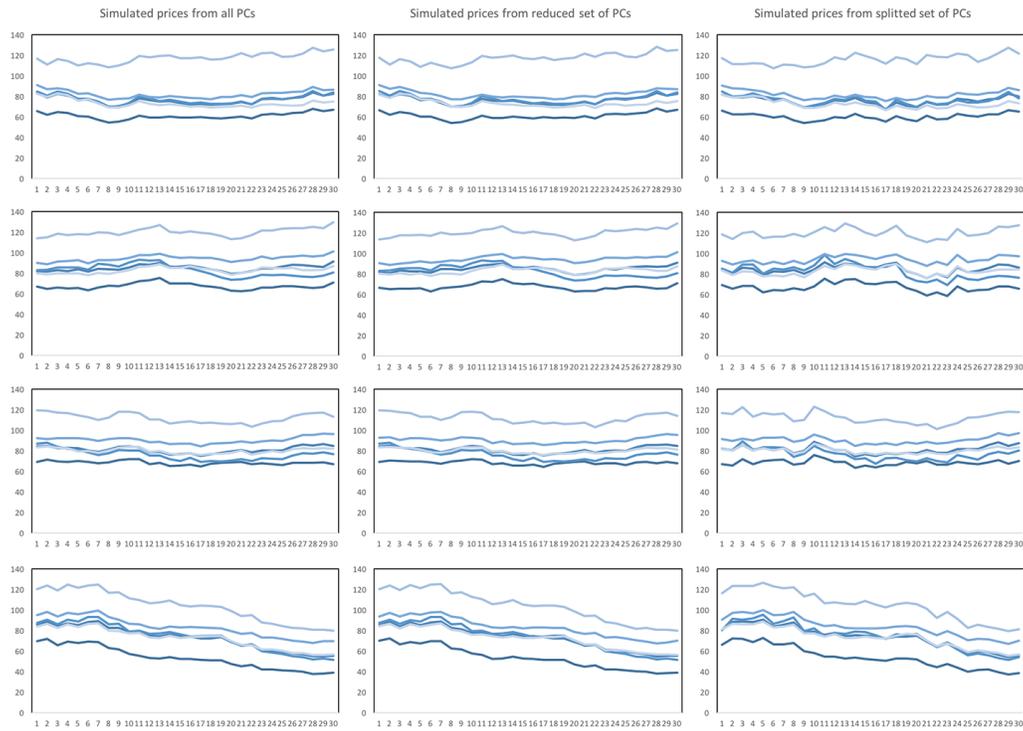


Figure 6: 4 simulations of correlated asset movements comparing simulated principal components

When simulating with GBM one draws random numbers from a standard normal cumulative distribution. This will result in an  $\tilde{\mathbf{X}}$  matrix of normally distributed returns. Thus, the principal components will be normally distributed. As a linear combination of independent normally distributed random variables also is normally distributed, every  $Q_{r_i}(q|PC_1, \dots, PC_m)$  from equation 11 will be normally distributed. This does not mean that the simulated marginal distributions are normally distributed. Remember  $Q_{r_i}(q|PC_1, \dots, PC_m)$  is only a data point in the mapping of the cumulative return distribution, hence the distribution itself is not necessarily normally distributed. However, if a different stochastic process is used in the generation of  $\tilde{\mathbf{X}}$ , one can make the principal components follow different distributions.

### 2.2.3 Construction of Future Return Distributions

When a set of principal components is simulated, the estimation of future cumulative return distributions can begin. All assets in the data set use these principal components in several linear combinations to get the data points which make up the asset's cumulative return distribution. The constants in these linear combinations are the  $\alpha$ s and  $\beta$ s estimated in the beginning of the algorithm. This procedure makes a crucial assumption that these constants also hold in the future. The linear combination in question is the factor model in equation 11. As mentioned in section 2.2.1 these scatter plots must be interpolated to complete the cumulative return distributions for all assets. For this step, the algorithm uses a shape-preserving piecewise cubic Hermite interpolation. To create pleasing cumulative return distributions,

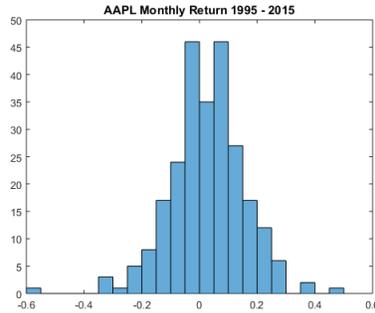


Figure 7: AAPL’s distribution of monthly returns from 1995 to 2015

two properties must be present.

- *Monotonicity*, as the function is a cumulative distribution, it should be strictly increasing that is  $\forall x, y$  were  $x \leq y, f(x) \leq f(y)$
- *Smoothness*, ideally the function should be as smooth as possible and not contain sharp bends.

Figure 5 shows how the algorithm draws a simulated return after the interpolation is done. However, this process does only result in one simulated return per asset. By simulating a new set of principal components and repeating this process many times, a return distribution will take form. Figure 8 shows how the monthly return distribution look like for 3 different stocks as the number of simulations increases. Notice how it takes at least 10,000 simulations before the distributions start to look continuous. For comparison AAPL’s monthly return from 1995 to 2015 is plotted in figure 7. The simulated monthly distributions clearly replicate the behaviour of a real stock.

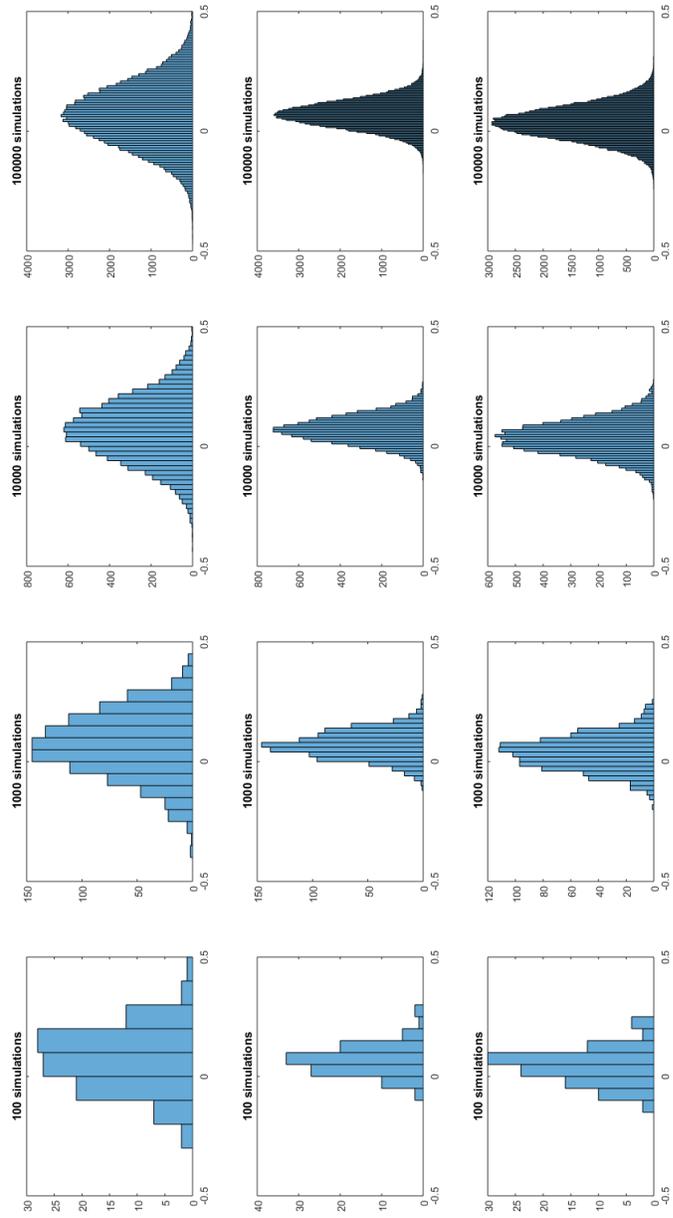


Figure 8: Development of return distributions as the number of simulations increases

## 2.3 Risk Parity Portfolio Optimization

With estimates of future return distributions at hand, one can use this information to build a portfolio in which each asset contributes equally to the total portfolio risk. Such a portfolio is called a *risk parity portfolio*, and the intuition behind this approach is to obtain the largest degree of diversification achievable, and maximize the long term risk adjusted return [13].

### 2.3.1 Mathematical Representation

To achieve risk parity, the total risk contribution (TRC) from each asset to the portfolio must be equal. TRC from an asset  $i$  is calculated using equation 30,

$$TRC_i = x_i MRC_i = x_i \frac{\partial R_p}{\partial x_i} \quad (30)$$

where  $MRC_i$  is defined as the marginal risk contribution from  $i$ ,  $R_p$  is the portfolio risk and  $x_i$  is the asset weight. The total risk of the portfolio is the sum of the individual risk contributions,

$$R_p = \sum_i TRC_i = \sum_i x_i MRC_i = \sum_i x_i \frac{\partial R_p}{\partial x_i} \quad (31)$$

The mathematical solution to a risk parity portfolio of  $n$  assets is a vector  $\mathbf{x}$  of  $n$  weights where the following equation holds,

$$x_1 \frac{\partial R_p}{\partial x_1} = \dots = x_n \frac{\partial R_p}{\partial x_n} \quad (32)$$

As there is no analytical solution to this problem, numerical optimization must be used. Equation 33 formulates the optimization problem where the

solution is a risk parity portfolio [14]. The target function to be minimized is the sum of all squared differences in total risk contribution from the assets. The last constraint is included to force long positions.

$$\begin{aligned}
& \underset{\mathbf{x}}{\text{minimize}} && \sum_i \sum_j (TRC_i - TRC_j)^2 \\
& \text{subject to} && \\
& && \sum_i x_i = 1 \\
& && 0 \leq x_i \leq 1, \forall i
\end{aligned} \tag{33}$$

Notice how these equations are not risk measure specific, indicating the flexibility mentioned in chapter 1. In the risk parity heuristic the investor may choose which risk measure to use as opposed to other heuristics where only volatility is used. A risk averse investor interested in his portfolio's tail behaviour, might prefer to use Conditional Value at Risk instead of volatility. With risk parity, this option is possible and the solution lies in the way the TRCs are being calculated.

### 2.3.2 The Optimization Algorithm

The optimization algorithm takes as input the matrix,  $\mathbf{D}$ , where

$$\mathbf{D} = \begin{bmatrix} r_{1,1} & \cdots & r_{1,n} \\ \vdots & \ddots & \vdots \\ r_{\tilde{s},1} & \cdots & r_{\tilde{s},n} \end{bmatrix} \tag{34}$$

containing  $\tilde{S}$  simulated returns for all  $n$  assets generated by the simulation algorithm in section 2.2. In this matrix, every column makes up the return

distribution for the given asset. Assuming one holds a portfolio,  $\mathbf{x}_p$ , where  $\mathbf{x}_p$  is a column vector of portfolio weights, the return of the portfolio for a given simulation will be given by equation 35

$$r_p = \sum_i r_i x_i \quad (35)$$

Applying the same equation on all simulated returns, one can obtain the portfolio return distribution. Equation 36 shows this step in matrix form,

$$\mathbf{D}_p = \mathbf{D}\mathbf{x}_p = \begin{bmatrix} r_{1,1} & \cdots & r_{1,n} \\ \vdots & \ddots & \vdots \\ r_{\bar{s},1} & \cdots & r_{\bar{s},n} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} r_1 \\ \vdots \\ r_{\bar{s}} \end{bmatrix} \quad (36)$$

which produces a return distribution vector for the portfolio consisting of the assets as specified by the weight vector  $\mathbf{x}_p$ .

With the portfolio return distribution one can calculate the value of a risk measure. For each iteration in the optimization algorithm, the risk of the portfolio for the current  $\mathbf{x}_p$  is calculated. Next a marginal change,  $\Delta x_i$ , is made to the weight of asset  $i$ , and a new portfolio risk is calculated using the new  $\mathbf{x}'_p$ . The resulting change in portfolio risk is used to calculate the current  $MRC_i$ .  $TRC_i$  is then found using equation 37<sup>3</sup>,

$$TRC_i = x_i \frac{\partial R_p}{\partial x_i} \approx x_i \frac{R_p(\mathbf{x}'_p) - R_p(\mathbf{x}_p)}{\Delta x_i} \quad (37)$$

The algorithm then calculates the target function value to be minimized, which is the sum of all squared differences in TRCs, stated in equation 33.

---

<sup>3</sup>Notice how the derivation in equation 37 is merely a numerical computer implementation of  $\frac{\partial f}{\partial x} = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$

After a number of iterations, a portfolio  $\mathbf{x}_p$  is found where the result satisfies equation 32 and the solution is returned.

As mentioned in section 2.3.1, this procedure is the same for every risk measure where  $TRC_i$  has to be calculated numerically. However, when volatility is chosen as risk measure, an analytical formula can be used. In this case equation 37 becomes

$$TRC_i = x_i \frac{\partial \sigma_p}{\partial x_i} = x_i \frac{\sum_j x_j \sigma_{ij}}{\sigma_p} \quad (38)$$

where  $\sigma_p$  is the portfolio volatility given by equation 39.

$$\sigma_p = \sqrt{\sum_i \sum_j x_i x_j \sigma_{ij}} \quad (39)$$

Here  $\sigma_{ij}$  is the covariance between asset  $i$  and  $j$ , and  $\sigma_{ii}$  is defined as the variance of asset  $i$ .

### 2.3.3 Risk Measures in Risk Parity Optimization

In risk parity research, volatility has been the risk measure receiving most attention. Even though several authors mention the possibility to use different risk measures as one of the major benefits to using risk parity [14], [15], an overwhelming majority of recent research papers focus on risk parity in comparison with other heuristics, using volatility as risk measure [13], [16], [17].

As volatility does not indicate the direction of an investment's movement it is not ideal for investors in practice. Investors are not distressed by gains, and volatility, treating all fluctuation as risk, will therefore perform poorly

for non-normal distributions. The important tail information is not captured by volatility and therefore it lacks focus on what matters most.

*Downside deviation* partly addresses these issues, as it focuses only on the negative fluctuation, ignoring the positive portfolio gains. As defined in equation 40, downside deviation calculates every return below some predefined threshold  $c$ , and reports this back as risk.

$$\text{Downside deviation} = \sqrt{\frac{1}{n} \sum_{r_i < c}^n (r_i - c)^2} \quad (40)$$

In this way, downside deviation estimates the potential losses of the portfolio to some degree. However, it does not offer any insight on the probabilities of the different levels of loss, and so it fails to become the preferred risk measure to be used by investors.

*Value at Risk* (VaR) focuses on the tail of the return distribution, and answers the question "*What is the most I can expect to lose with a given confidence interval over a time period?*". Figure 9 shows VaR as the breaking point for the "worst case scenario" - defined as the highest percentage of portfolio value the investor may expect to lose with  $1 - \alpha\%$  certainty over the time period considered.

Even though VaR has been widely used for measuring risk in the financial industry, it still has severe shortcomings when used alone as a risk measure. VaR does define the break point for the  $\alpha\%$  worst case loss, but completely ignores the size of the expected loss once this point is breached. This is where CVaR comes in. Defined as in equation 41 and shown in figure 9, CVaR measures the expected loss of the portfolio in the  $\alpha\%$  worst cases.

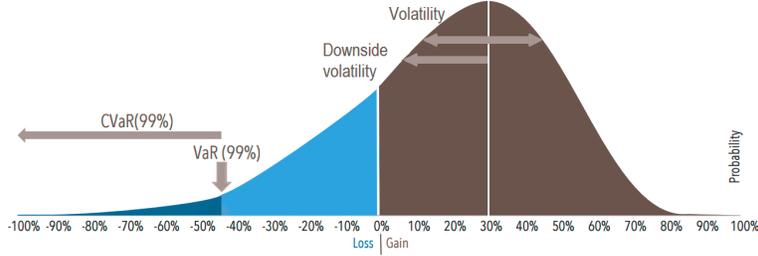


Figure 9: Risk measures discussed

$$CVaR_{1-\alpha}(r_i) = \int_{-\infty}^{VaR_{1-\alpha}} r_i p(r_i) dr_i \quad (41)$$

Here  $r_i$  and  $p(r_i)$  is asset  $i$ 's stochastic return and its probability distribution, respectively.

CVaR is more sensitive to the shape of the tail distribution and it estimates the risk of an investment in a more conservative way, focusing on the least profitable outcomes. As VaR has shortcomings when used as a risk measure, it also has mathematical properties far inferior to those of CVaR. Rockafellar & Uryasev [18] show how VaR, when calculated from scenarios, is a non-convex and discontinuous function, making it difficult or even impossible to optimize as it has many local extrema. In this framework, where the input are discrete distributions calculated from scenarios, optimization using VaR as risk measure is not performed.

The main objective of the simulation algorithm presented in section 2.2 is to simulate marginal return distributions preserving asset specific tail behaviour. This information is of interest to CVaR as risk measure and when using CVaR as risk measure in risk parity optimization. Given CVaR's superior mathematical properties as compared with those of VaR, and given

its characteristics as a risk measure used in optimization, CVaR will be the risk measure of focus in the empirical study later in this thesis.



## 3 Empirical Results

This chapter presents the results from testing the simulation algorithm on an empirical data set. Without any knowledge of how the stock markets moved in January 2016, the algorithm will use available historical data and simulate the performance of stocks in January using the framework presented in chapter 2. A brief introduction to the data set is given before section 3.2 to section 3.5 describe the realized factor models and simulation results. Next a performance test of the simulation algorithm is presented in section 3.6 where the simulated returns are compared to the actual realized returns of January 2016. The results of the risk parity optimization are presented in section 3.7.

### 3.1 Data Description

The data set in this empirical test consists of 80 stocks from the S&P100 index. The S&P100 index is an index composed of a subset of companies from Standard&Poor's 500 index. The companies listed in the S&P100 index are mostly the leading well-established publicly traded U.S companies, representing a large amount of the total market capitalization of the U.S equity markets. In order to get higher effect of PCA, this 80 stock data set is split into 10 smaller data sets with respect to the Global Industry Classification Standard (GICS). Descriptive statistics of historical data for all stocks grouped into GICS sectors are listed in appendix C.

Monthly adjusted close prices taken from January 1st 1995 to December 31st 2015 output 252 data points for each stock. Appendix B shows a list

of all the stocks in the data set. These time series have been converted to a return matrix with discrete returns calculated using equation 42,

$$r_t = \frac{S_t - S_{t-1}}{S_{t-1}} \quad (42)$$

implying that one time step equals one month. The reason why 20 stocks are excluded from the data set is because these companies do not have data going back to 1995, as the majority is younger than 20 years old.

One month is a reasonable period to hold a portfolio of assets before rebalancing and was therefore seen as a natural choice for the simulation period resulting in a  $\Delta t$  value equal to 1 with months as unit. The number of simulation steps into the future is also set to 1 in the algorithm, meaning the results at the end of the run are stock returns in January 2016. As figure 8 points out, several thousand simulations is necessary to produce continuous return distributions. In this test, 100.000 simulations for each stock were used, meaning  $\tilde{S}$  equals 100.000.

When it comes to the decision of what historical data to use, there are many options to choose from and no optimal answer. A key element to consider is what applications the simulation results are going to be used in. Thus, the choice of daily, weekly or monthly data should reflect the intended use. The time horizon of the historical data should also be considered. Too few data points and the effect of quantile regression will become absent. With too many data points, the effect of one data point would be diluted resulting in a loss of stock specific tail behaviour.

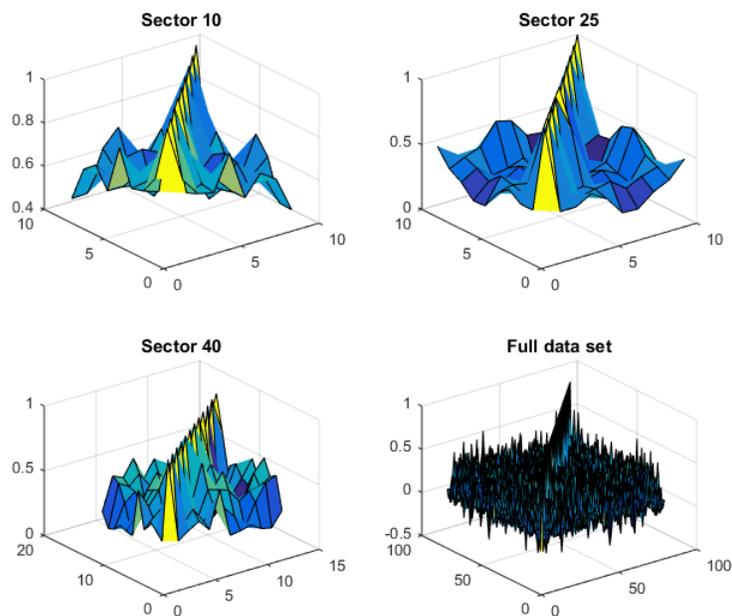


Figure 10: Comparison of correlation matrices between data sets

## 3.2 Empirical Model

The simulation algorithm is run on each GICS sector individually, meaning every sector with its respective stocks is treated as an independent data set. This is done to achieve a higher correlation in the data sets, thus enhancing the effect of using PCA in the simulations. Figure 10 shows a comparison of the correlation matrix of the historical data between the arbitrarily selected sectors 10, 25 and 40, and the full data set <sup>4</sup>. These surface plots show an increased internal correlation when splitting the full data set into GICS sectors indicating, as expected, an increased effect of PCA.

<sup>4</sup>See appendix D.2 for a comparison between all sectors and the full data set.

Sector	Columns	Columns After PCA	Reduction
10	9	7	22,2%
15	2	2	0,0%
20	13	11	15,4%
25	10	8	20,0%
30	9	8	11,1%
35	12	10	16,7%
40	12	9	25,0%
45	9	8	11,1%
50	2	2	0,0%
55	2	2	0,0%
All	80	50	37,5%

Table 1: Effect of PCA with at least 95% variance explained

Table 1 shows the effect of using PCA keeping at least 95% of the variance in all data sets. The third column in the table shows how many risk factors are included in the PCA factor models, representing the variable  $m$  from chapter 2. 95% is used as a variance margin to achieve dimension reduction and at the same time keep individual stock characteristics in the data sets. The fourth column shows the percentage effect of using PCA in the model to reduce complexity. If there was any reduction, over 10% was achieved. With a variance margin below 80% only common trends will be kept in the data set after PCA, making the simulation of stock specific tail characteristics impossible.

Figure 11 shows a plot of the cumulative variance explained and how much

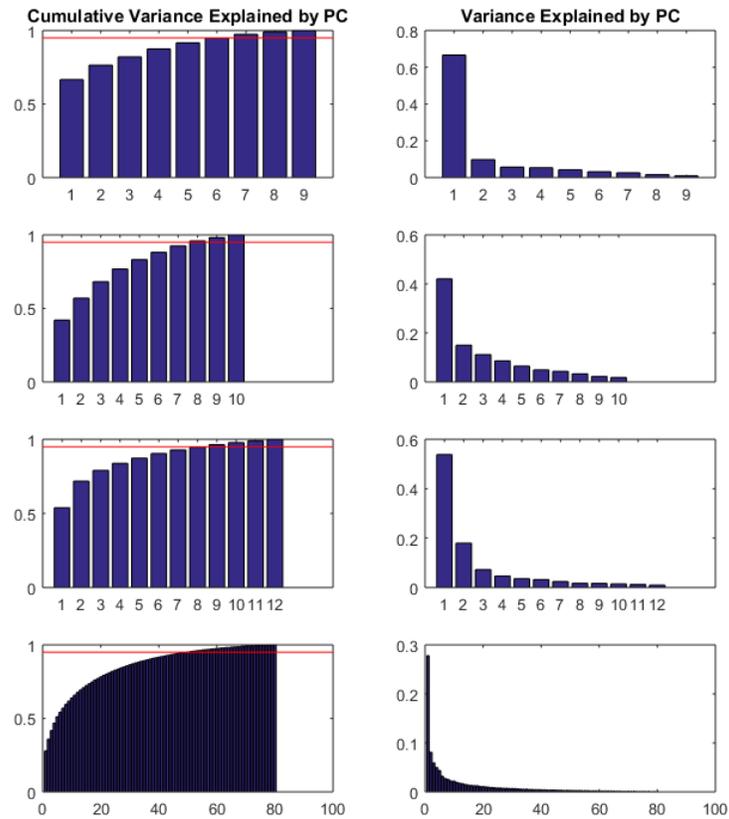


Figure 11: Effect of PCA, comparison between sector 10, 25, 40 and the full data set

variance each principal component explains for sectors 10, 25, 40 and the full data set <sup>5</sup>. The red horizontal line in the first column indicates the 95% mark. With equation 8 in mind, notice how small the eigen values of the principal components in the full data set are. From the twentieth principal component and onwards, each PC explains less than 1% of the variance in the data set. Using the full data set of 80 stocks and keeping at least 95% variance in the simulation would require 50 principal components in the PCA factor models. With that many risk factors, the algorithm ran into problems with numerical instability in the simulation of future cumulative return distributions. By splitting the data set into GICS sectors, each factor model had fewer PCs which eliminated the risk of overfitting in the interpolation caused by the instability. Figure 12 illustrates this instability problem. Remember that every data point is normally distributed as discussed in section 2.2.2. Notice how the means create a smooth S-shaped cumulative distribution. With too large variance in the data points, they may not be aligned properly, thus creating an undesired cumulative return distribution under interpolation <sup>6</sup>. This shows how high positive correlation in the data set is important for the PCA and the stability in the framework.

### 3.3 Quantile Regression Results

After the PCA is run on the historical data, the principal components are ready to be used in the quantile regressions. This is the second process in the

---

<sup>5</sup>See appendix D.1. for a comparison between all sectors and the full data set

<sup>6</sup>See appendix E for an explanation of why the variance in data points grows with the number of principal components in the factor model

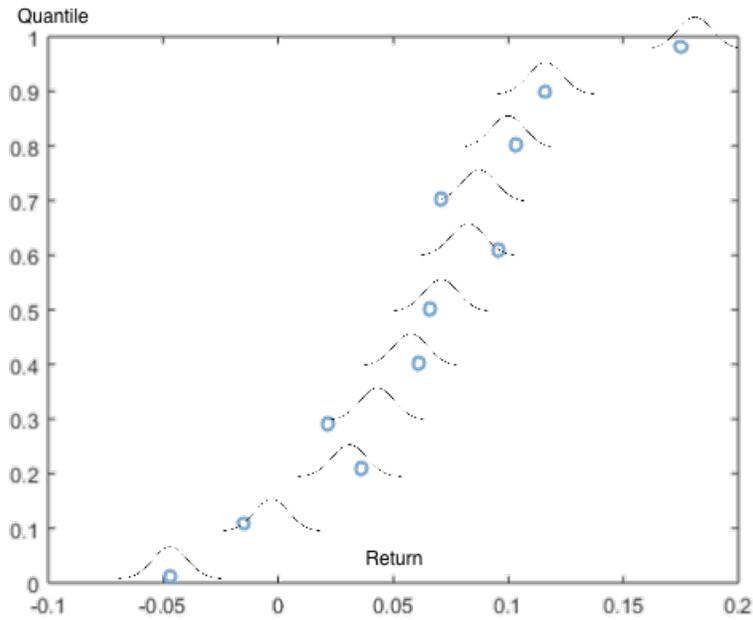


Figure 12: The cumulative distribution with too many explanatory variables

flow chart in figure 4. The number of explanatory variables from equation 11 used in the regressions, are shown in the third column in table 1. The  $\beta$  is as mentioned in section 2.1, sensitivity to the risk factor, while the  $\alpha$  measure the risk less behaviour, that is the stock's trend. Figure 13 shows the  $\alpha$  and 3  $\beta$ s for APA, XOM, ABT and LLY's stock. Remember  $\beta_1$  is the stock's sensitivity to the principal component that explains the most variance, then comes  $\beta_2$  and so on.

Because variance is explained by the principal components, the  $\alpha$  plot indicates the cumulative probability distribution for the stock's trend based on 252 data points of historical monthly data. APA's 0.5 quantile in the  $\alpha$  plot indicate a median return of 1,11% per month which corresponds well with its mean return of 1,15% found in the descriptive statistics in appendix

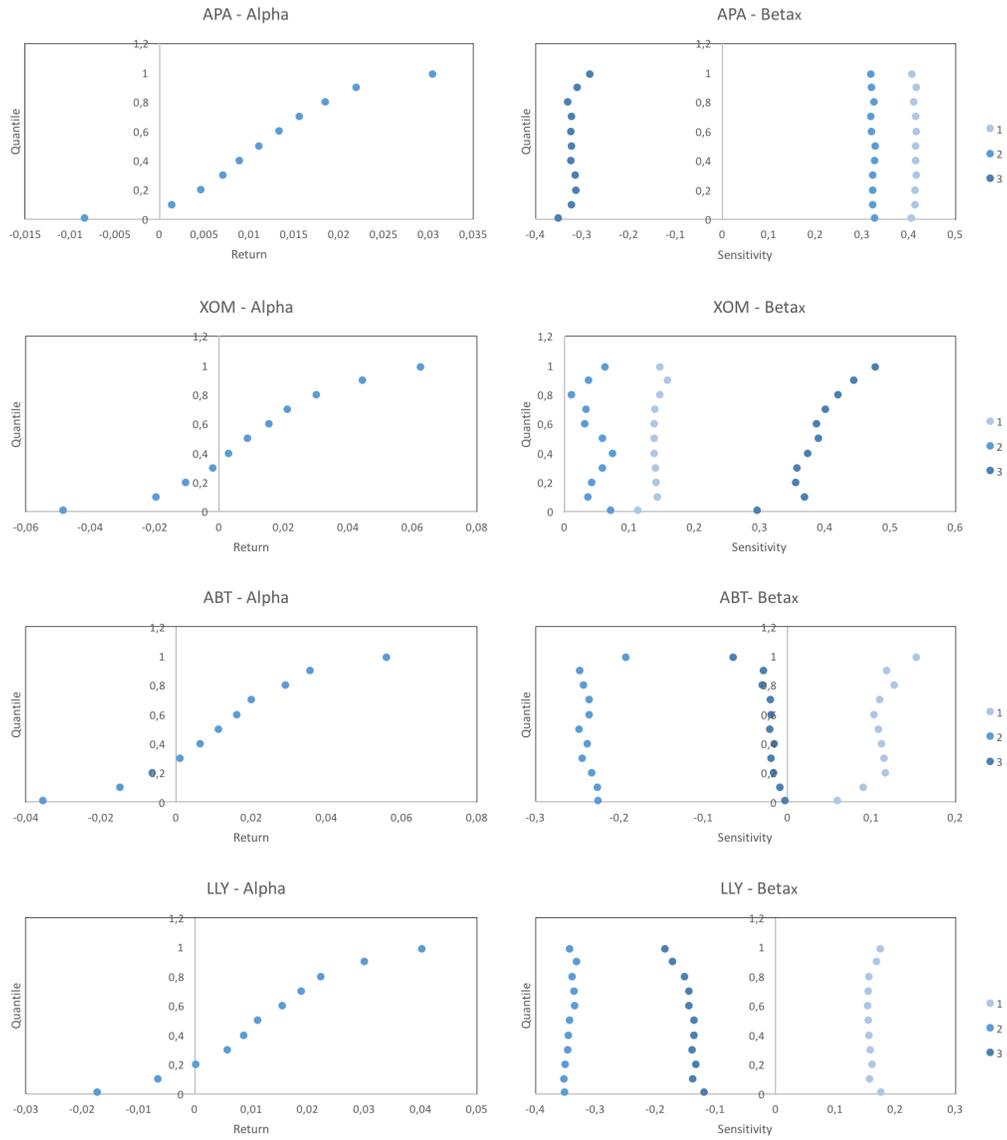


Figure 13:  $\alpha$  and  $\beta$  estimates from quantile regression

C. If the regressions were based on daily data, one would expect the median return to be closer to 0.

Notice how the  $\beta$  plots reveal interesting characteristics of the return distribution before the distribution is simulated. APA and LLY have almost no difference in sensitivities when comparing between quantiles as opposed to XOM and ABT. A typical I-shaped *beta* plot will reflect in smaller differences between each quantile making the effect of quantile regression lower. Thus, a sharper simulated cumulative return distribution is expected. This assumption will be investigated further in section 3.4.

A second characteristic one can detect by looking at the  $\beta$  plots is an indication of the stock's volatility. Higher absolute  $\beta$  values make the stock more sensitive to risk factors, which will be reflected in the stock's behaviour. APA and LLY have higher absolute  $\beta$  values than XOM and ABT meaning APA and LLY should be more volatile. By comparing volatility in the descriptive statistics, this claim can be confirmed. Studying the simulated return distributions in appendix G, one can clearly see that this is also the case in the simulated future.

### **3.4 Simulated Cumulative Return Distributions**

Quantile regression makes it possible to create a future cumulative return distribution presented in section 2.2.3. However, the effect of this regression varied in the test. Figure 14 shows 10 simulated cumulative return distributions for APA, XOM, ABT and LLY. From the discussion in section 3.3 one would expect I-shaped cumulative return distributions for APA and LLY because of the small differences in  $\beta$  values between quantiles. On the other

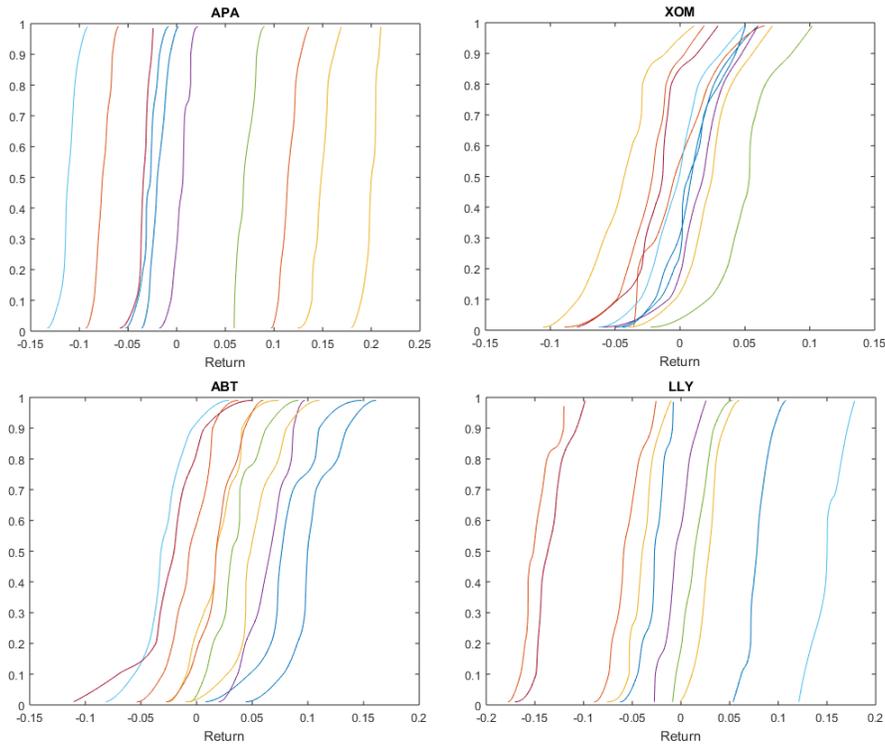


Figure 14: Simulated cumulative stock return distributions for APA, XOM, ABT and LLY

hand, XOM and ABT show more desired patterns where the variation is bigger leading to S-shaped cumulative distributions.

Notice how the vertical shifts are greater for APA and LLY than they are for XOM and ABT. This behaviour can also be predicted by looking at the plots of the  $\beta$  values. The stocks with greater vertical shifts have higher absolute  $\beta$  values than the stocks with compact cumulative distributions. The mechanics behind the relationship between the plots in figure 13 and figure 14 lie in the factor model in equation 11.

One of the reasons why some stocks have I-shaped cumulative return

distributions lies in the historical data. With clustered data points the regression lines illustrated in figure 1 will lie close together which will lead to a lower effect of quantile regression. Using monthly data increases the possibility of this outcome in two ways. Firstly, 252 data points might not always be enough to clearly separate the regression lines for different quantiles. Secondly, 252 monthly data points might not capture enough variance. If some share prices increase in 9 out of 10 months, there will be a clustering in the data. With daily data there are roughly 30 times as many data points per month than with monthly data, making the risk of clustering lower.

### 3.5 Simulated Return Distributions

The result of running the simulation algorithm will be as mentioned a matrix **D**, consisting of simulated stock returns. These distributions are simulated as a correlated set of stocks within each GICS sector. Using 100,000 simulations the negative effect of I-shaped cumulative distributions will be reduced. In addition, the vertical shifts make the distribution keep important tail information as stock returns will be drawn from the entire return interval.

Figure 15 shows the probability distributions for four different stocks. Notice how different the characteristics of each distribution are. By gathering individual stock characteristics from empirical data, the algorithm can predict the future distribution keeping the trends and tail behaviour. Table 2 shows the descriptive statistics for these simulated distributions. The skewness for all stocks is slightly positive meaning the right tale is longer than the left, but the low skewness values along with the Jarque-Bera test <sup>7</sup> results

---

<sup>7</sup>Jarque-Bera test, 1 if the null hypothesis is rejected at a 5% significance level, 0

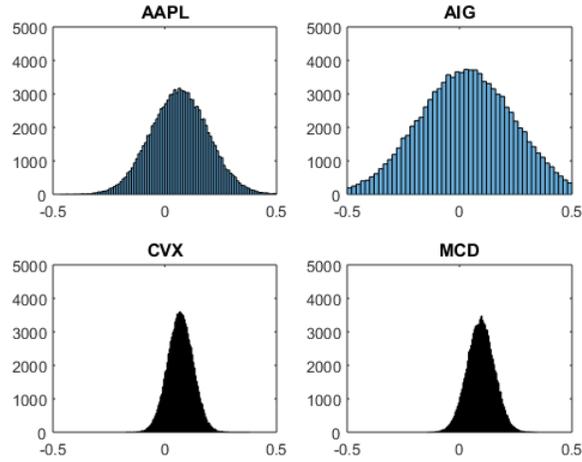


Figure 15: Simulated stock return distributions for January 2016

Stock	Mean	St.Dev	Skewness	Kurtosis	JB Test
AAPL	0.066241	0.12819	0.017501	2.9905	0
AIG	0.031098	0.2138	0.008025	2.9854	0
CVX	0.069666	0.056197	0.002075	3.0297	0
MCD	0.094022	0.059688	0.005722	2.997	0

Table 2: Descriptive statistics of probability distributions from figure 15

indicates a plausibility that the distributions are normal distributions. Also the kurtosis values support this claim.

Comparing the new descriptive statistics in table 2 with the statistics in appendix C one can see a noticeable difference. Especially AIG with a historic skewness and kurtosis value of 6.52 and 73.89 respectively. The reason why these values differ, is because the simulated principal components are otherwise.



Figure 16: Monthly development of the S&P500 index from 1995 to Q2 2016.

normally distributed. The return distributions will therefore follow a normal distribution closely as the new statistics indicate. In addition, the historical statistics are an average of the past, calculated using 252 returns versus 100.000 in the simulation. Given AIG's turbulent past, such extreme statistics are not surprising and differences between new and historical statistics may occur for some stocks.

Figure 24 in appendix G show the simulation results for all stocks. Some stocks have higher volatility than others and some are more skewed than others. What the majority have in common is that almost all means are slightly positive, meaning the algorithm expects most stocks to make money in January 2016. Looking at the monthly development of the S&P500 index in figure 16, this assumption is not entirely wrong as the clear trend is a positive monthly return.

### 3.6 Simulation Performance

To measure the performance of the simulation algorithm, different confidence intervals have been calculated from the simulated distributions to check if the realized returns were within the intervals. These intervals are the 60, 70, 80, 90 and 95% confidence intervals. Table 3 shows an analysis of the confidence intervals for the four stocks from section 3.5. Column two shows the realized stock return in January 2016 and the other columns investigate if the realized return lies inside its respective interval. Ideally, 95% of the times the realized stock returns should lie in the 95% confidence interval predicted by the simulation algorithm. Using the data in table 3, one can produce a price plot of the stock including both historical development and a forecast with different certainty levels. Examples of such plots are illustrated in figure 17.

Table 4 shows the amount and relative amount of stocks, where the realized January return was in the interval. This table is a good indication on how well the algorithm performed in simulating mean returns with the 80 stocks from the S&P 100 index in January 2016. Notice how less than 30% of the returns was in the 60% confidence interval, while almost 80% of the returns was inside the 95% confidence interval. Because the algorithm gathers stock characteristics from historical data, naturally the algorithm will perform better in times where the markets behave similar as the trends in the past. As January 2016 suffered from turbulent markets and most of these companies lost money, one can not say the month was an average month when the historical period in question is January 1995 to December 2015. Figure 18 shows data from table 3 included mean of the four stocks

Ticker	JanuaryReturn	Lower60	Upper60	Inf60Interval	Lower70	Upper70	In70Interval	Lower80	Upper80	Ins80Interval	Lower90	Upper90	In90Interval	Lower95	Upper95	Ins95Interval
AAPL	-7.52 %	-4.23 %	17.42 %	False	-6.74 %	19.87 %	False	-9.87 %	23.00 %	True	-14.34 %	27.77 %	True	-18.36 %	31.84 %	True
AIG	-8.86 %	-14.92 %	21.17 %	True	-19.13 %	25.33 %	True	-24.35 %	30.53 %	True	-31.99 %	38.23 %	True	-38.74 %	44.74 %	True
CVX	-3.88 %	2.25 %	11.71 %	False	1.16 %	12.79 %	False	-0.21 %	14.16 %	False	-2.29 %	16.21 %	False	-4.09 %	17.91 %	True
MCD	4.78 %	4.38 %	14.43 %	True	3.23 %	15.62 %	True	1.75 %	17.09 %	True	-0.42 %	19.23 %	True	-2.31 %	21.11 %	True

Table 3: Confidence intervals of probability distributions from figure 15

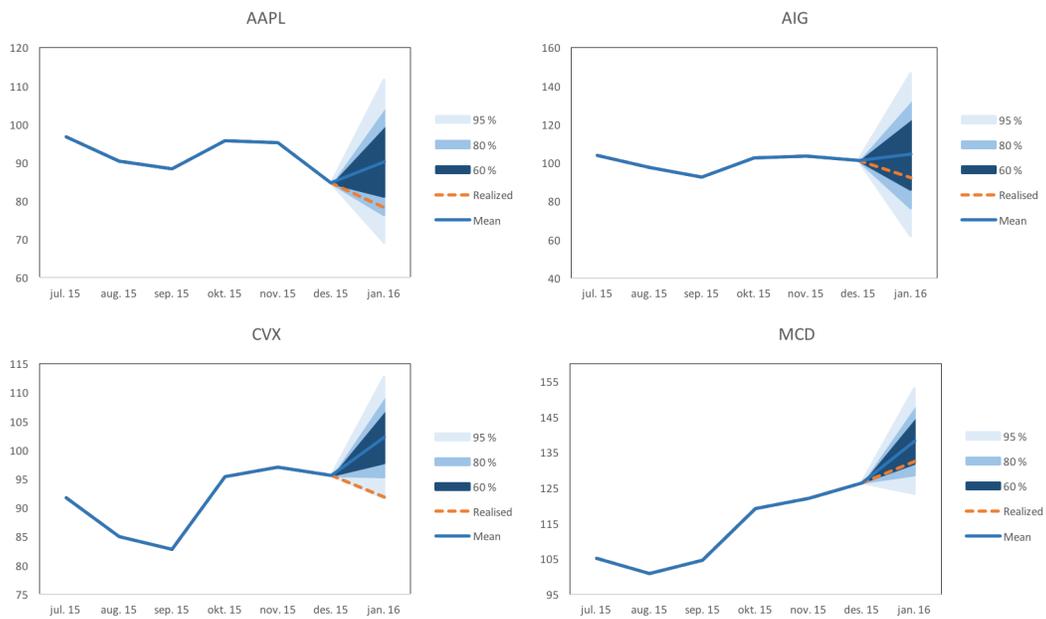


Figure 17: Forecast plot for January 2016 using probability distributions from figure 15

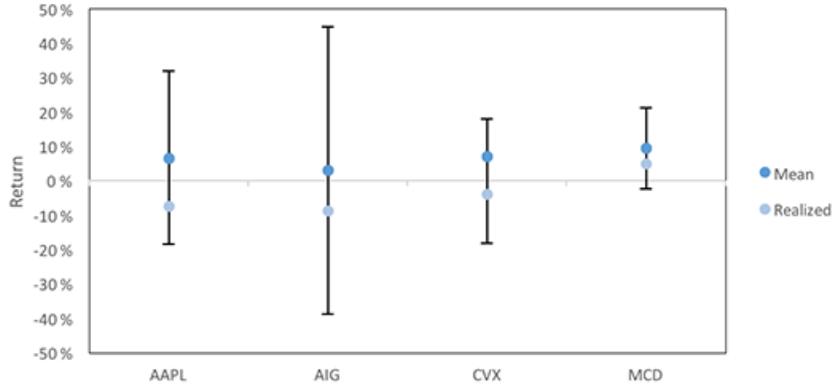


Figure 18: Realized and mean returns with 95% confidence intervals

	In 60%Int	In 70%Int	In 80%Int	In 90%Int	In 95%Int
Amount	23	28	41	51	63
Relative	29%	35%	51%	64%	79 %

Table 4: Relative amount in CI Intervals

along with the 95% confidence intervals and the realized return <sup>8</sup>.

By looking at the performance results of the simulation algorithm in January 2016 comparing the estimated mean returns to the realized returns, one might conclude that the algorithm performed badly. However, this is not enough to discard the framework even though the results could be better. Various periods of historical data, forecasting lengths and time intervals are all variables that should be changed and tested with the algorithm before one can say with great certainty that the framework does not work. That type of analysis is beyond the scope of this thesis. This chapter is, as mentioned in chapter 1, merely about demonstrating how one uses the algorithm to

<sup>8</sup>See appendix F for a similar plot of all stocks.

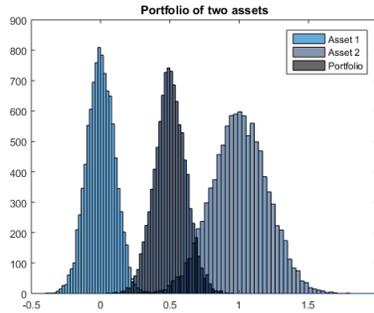


Figure 19: Example portfolio of two stocks with different characteristics.

simulate real life stocks and what data one can extract from the  $\mathbf{D}$  matrix the algorithm produces. In addition, with regards to the portfolio optimization model, the goal of the simulation algorithm is not necessarily to estimate precise mean returns, but rather model the risk characteristics as accurately as possible.

### 3.7 Portfolio Optimization

The portfolio optimization process uses the algorithm presented in section 2.3.2, with  $\mathbf{D}$  as input. This matrix is then split up into GICS sectors and the optimization algorithm is run on every sector. This is done to force higher correlation in the data sets, as it can be shown that for CVaR stocks from different sectors may have too different characteristics, making it hard or even impossible to achieve risk parity for any chosen set of weights.

For example, figure 19 presents the return distributions of two stocks from different sectors. Their volatility is always positive, implying that there will always exist a set of portfolio weights which equalizes the total risk contribution from the two stocks in a given portfolio. However, with CVaR

as risk measure, a portfolio consisting of the two stocks will have marginal risk contributions with unequal sign. Stock 1 will always contribute to an increase in portfolio risk, as long as  $x_2 > 0$ , and stock 2 will contribute to a decrease in portfolio risk, as long as  $x_1 > 0$ , in other words,

$$\frac{\partial R_p}{\partial x_1} < 0 < \frac{\partial R_p}{\partial x_2} \quad (43)$$

and equation 32 may never hold for any vector  $\mathbf{x}$  of non-negative weights.

After the data set is split up into individual sectors, the stocks in each subset share more of the same characteristics. This implies a higher degree of correlation and ultimately distributions with less extreme differences than those shown in the example in figure 19. This is also observable in the correlation matrices in appendix D.2, where the correlation matrix for the full data set contain both positive and negative values, whereas the correlations in the individual sectors span values between zero and one.

As the stocks are split up into GICS sectors, the problem of uncorrelated assets is solved and risk parity optimization using CVaR may be performed.

### 3.7.1 Optimization Results

The optimization framework is implemented using three different risk measures; CVaR, volatility and downside deviation, where the last two are included for comparison. The risk calculations described in section 2.3.2 are implemented for the respective risk measures, and the target function for each risk measure is minimized, observing the restrictions. The resulting weight vector  $\mathbf{x}$  is then returned along with a vector of the associated TRCs to check that equation 32 is indeed satisfied for the given weights.

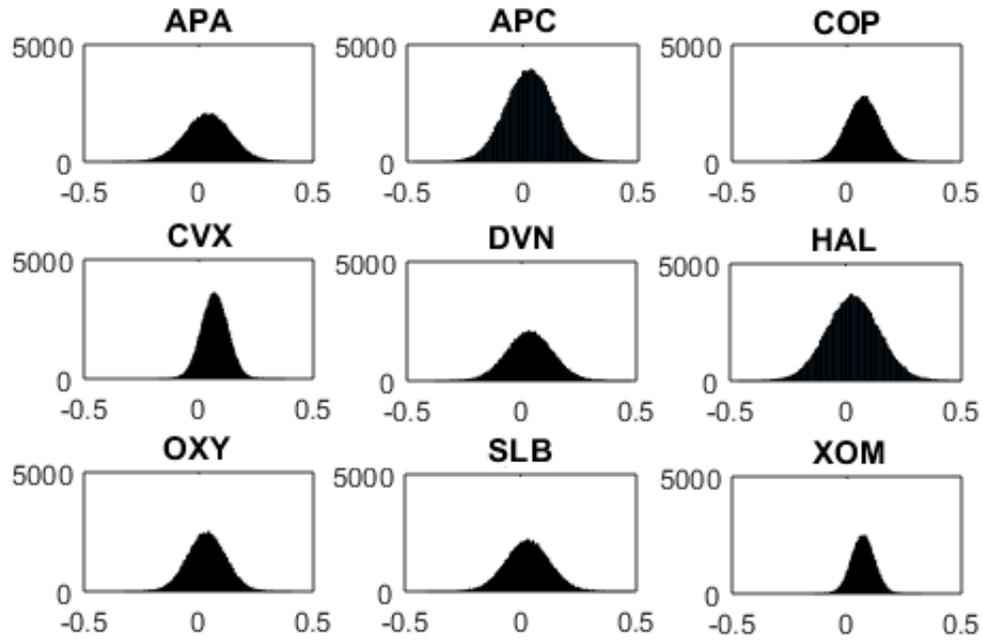
<b>Sector</b>	<b>CVaR</b>	<b>Volatility</b>	<b>Down dev</b>
<b>10</b>	2,47E-10	3,46E-12	3,85E-08
<b>15</b>	8,48E-14	5,91E-15	7,52E-08
<b>20</b>	5,36E-12	1,24E-12	9,77E-09
<b>25</b>	2,15E-07	8,90E-12	4,57E-08
<b>30</b>	4,99E-10	1,06E-12	1,28E-08
<b>35</b>	4,59E-07	1,09E-11	9,89E-08
<b>40</b>	5,36E-08	4,70E-12	3,10E-08
<b>45</b>	3,85E-13	7,60E-14	2,14E-08
<b>50</b>	9,45E-14	6,58E-19	2,79E-11
<b>55</b>	3,16E-12	6,93E-16	5,64E-07

Table 5: Target function values for the different risk measures used in the portfolio optimization

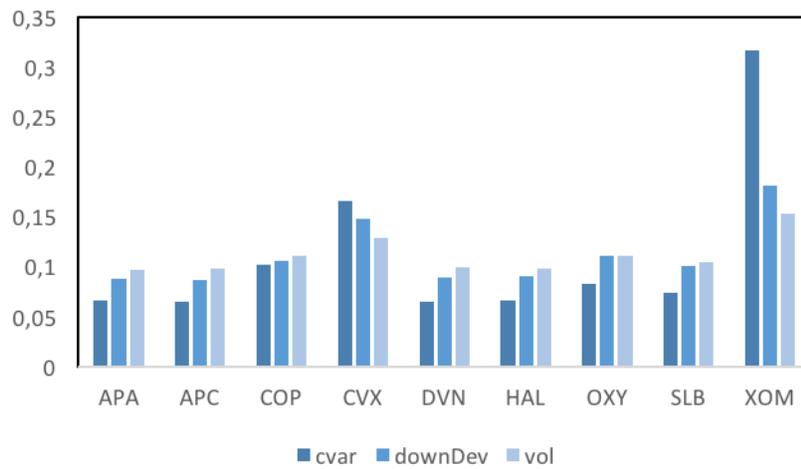
The target function values for each optimization run are presented in table 5. For each of the different sectors, all values approach zero, indicating risk parity.<sup>9</sup>

When observing the output of the optimization for the different sectors, it becomes apparent how the choice of risk measure is non-trivial when the simulation framework described in this thesis is used. In figure 20 the simulated distributions for the stocks in sector 10 are presented along with the weights allocated using the different risk measures. For volatility and downside deviation, the allocations are quite evenly distributed, as these consider only the spread of the distributions, ignoring tail behaviour. The allocations

<sup>9</sup>The risk parity weights for all sectors and risk measures are presented in appendix H.



(a) Simulated distributions for stocks in sector 10



(b) Allocated risk parity weights for sector 10

Figure 20: Simulated distributions and risk parity weights for sector 10

using CVaR, on the other hand, differ much more, allocating almost twice the weight to XOM as compared with that of the runner up, CVX. The first two risk measures also allocate more weight to these two stocks compared with the rest because the spread of the distributions for these two stocks is smaller. As CVaR focuses on the information in the tail end of the distributions, it is better able to capture the unique characteristics of the individual stocks which is shown clearly in this example.

Equivalently, the distributions of APA, APC, DVN and HAL have the longest tails on the negative side of zero, resulting in CVaR allocating the lowest weights to these stocks. Apparently, the information captured by the simulation algorithm is preserved, and the different characteristics of the risk measures become important.

The results of the optimization show that the simulated distributions may be relevant in portfolio optimization when using a practical risk measure like CVaR. There may be difficulties performing risk parity optimization with CVaR for large data sets of uncorrelated assets, but for smaller sets with higher correlation, this is not a problem. The unique characteristics of the different stocks are captured and taken into account to a larger extent when using CVaR as risk measure, compared to volatility and downside deviation.

## 4 Conclusion

This thesis presents a framework for simulating multivariate return distributions for correlated assets and a way to use these distributions to obtain a risk parity portfolio using a relevant risk measure like CVaR.

The solution overcomes some of the difficulties of earlier portfolio optimization models as the framework simulates the entire distributions and not just the conditional means. In addition the risk parity model does not rely on these means, but focuses only on risk as it examines the characteristics in the tails of the assets' distributions. This way, the input sensitivity is lowered and the model risk is decreased when used with CVaR, enabling better and more robust solutions to the portfolio optimization problem.

The marginal distributions generated in the simulation show how the independent behaviour and risk characteristics of the assets are preserved into the simulated future. The results of the portfolio optimization indicate how these characteristics result in different weight allocations for different risk measures, where CVaR captures the information in the tails of the distributions and allocates accordingly.

The optimization algorithm also shows how optimizing a large data set of relatively uncorrelated assets is a challenge using CVaR as risk measure. However, for smaller sets of assets with similar characteristics the algorithm finds optimal weight allocations with minimal differences in risk contributions. This shows how the framework can be used creating risk parity optimal solutions depending on tail information in the simulated distributions.

As the framework was tested only in one time period, an interesting next step would be to test the framework in different time periods to get a better

grasp of the general performance of the algorithm. The simulation and time step values are variables which could be changed, for example with daily returns and a one day simulation period. In this way one could get many more data points from the historical data, making the effect of the quantile regression higher.

The most interesting extension of the simulation algorithm would be to modify the simulation of future asset movements. By incorporating local volatility [19] with GBM one could obtain a dynamic volatility with a better fit to the market situation. Another possible extension is to use a different stochastic process than GBM to model heavier skewness and tails in the resulting principal component distributions. This would in the end impact the shape of the return distributions.

Further research may also be done on using the simulation output matrix in other applications than the portfolio optimization model implemented here. Other applications such as risk analysis, trading strategies and derivatives pricing may be relevant for the simulated distributions.

## References

- [1] Robert C. Merton (1980), *On Estimating Expected Return on the Market*, Journal of Financial Economics, Vol. 8, pp. 323-361
- [2] Harry Markowitz (1952), *Portfolio Selection*, The Journal of Finance, Vol. 7, No. 1, pp. 77-91
- [3] Robert D. Arnott, Jason Hsu & Philip Moore (2005), *Fundamental Indexation*, Financial Analysts Journal, Vol. 61, No. 2, pp. 83-99
- [4] Norges Bank Investment Management (2012), *Alternatives to Market-Value-Weighted Index*, NBIM Memo Equity Investments, Vol. 7
- [5] Carol Alexander & Melchior Ruppert (2013), *Factor Model Quantile Simulation of Stock Returns and Application to VaR and ETL Estimation*, Unpublished
- [6] Sébastien Maillard, Thierry Roncalli & Jérôme Teïletche (2010), *On the Properties of the Equally Weighted Risk Contributions Portfolios*, Journal of Portfolio Management, Vol. 36, No. 4, pp. 66-70
- [7] Denis Chaves, Jason Hsu, Feifei Li, & Omid Shakernia (2011), *Risk Parity Portfolios Vs. Other Asset Allocation Heuristic Portfolios*, Journal of Investing, Vol. 20, No. 1, pp. 108-118
- [8] William F. Sharpe (1964), *Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk*, The Journal of Finance, Vol. 19, No. 3, pp. 425-442

- [9] John Lintner (1965), *The Valuation Of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets*, The Review of Economics and Statistics, Vol. 47, No. 1, pp. 13-37
- [10] Stephen Ross (1976), *The Arbitrage Theory of Capital Asset Pricing*, Journal of Economic Theory, Vol. 13, pp. 341-360
- [11] Robert Koenker & Gilbert Basset Jr (1978), *Regression Quantiles*, Econometrica, Vol. 46, No. 1, pp. 33-50
- [12] Carol Alexander (2008), *Market Risk Analysis, Vol. 1 - Quantitative Methods in Finance*, John Wiley & Sons Ltd, ISBN13: 9780470998007
- [13] Edward Qian (2011), *Risk Parity and Diversification*, Journal of Investing, Vol. 20, No. 1, pp. 119-127
- [14] Denis Chaves, Jason Hsu, Feifei Li, & Omid Shakernia (2012), *Efficient Algorithms for Computing Risk Parity Portfolio weights*, Journal of Investing, Vol. 21, No. 3, pp. 150-163
- [15] Hossain Kazemi (2012), *An Introduction to Risk Parity*, Alternative Investment Analyst Review, Vol. 1, No. 1
- [16] Gregg S. Fisher, Philip Z. Maymin & Zakhar G. Maymin, *Risk Parity Optimality*, Journal of Portfolio Management, (Forthcoming)
- [17] Clifford S. Asness, Andrea Frazzini & Lasse H. Pedersen (2012), *Leverage Aversion and Risk Parity*, Financial Analysts Journal, Vol. 68, No. 1
- [18] Ralph Tyrrell Rockafellar & Stanislav Uryasev (2000), *Optimization of Conditional Value-at-Risk*, Journal of Risk, Vol. 2, No. 3

[19] Bruno Dupire (1994), *Pricing with a Smile*, Risk, Vol. 7, No. 1, pp. 18-20



## A Computational Information

The simulation algorithm can be implemented in any programming language. For the test in this thesis, it was implemented in *Matlab* which is a great tool for small to medium scale problems and enhanced for matrix algebra. There are also many toolboxes available which can simplify processes and reduce the amount of necessary code. Toolboxes used in the implementation are

- *Datafeed* Toolbox, used to download historical stock data from Yahoo! Finance.
- *Parallel Computing* Toolbox, used to simplify the coding for making the simulation loop run in parallel.

The program can run in parallel on a computer with several cores, or on a cluster network to significantly reduce the simulation time. The simulation program was tested on a PC running *Windows 7* with an 8 core processor and 16 GB of RAM. This simulation run time for each sector with 100.000 simulations, was less than 100 seconds.

For the quantile regressions, a script in *R* was written using the *quantreg* package. This script produces .csv files which are later imported by the program in *Matlab*. In addition, small scripts for preprocessing and formatting of data was written in the *Python* programming language.

All figures are created in *Matlab* or *Excel*.

## B Data set

### Data set

Ticker	Name	GICS Sector	Sector Code
APA	Apache Corp.	Energy	10
APC	Anadarko Petroleum Corp.	Energy	10
COP	ConocoPhillips	Energy	10
CVX	Chevron	Energy	10
DVN	Devon Energy	Energy	10
HAL	Halliburton	Energy	10
OXY	Occidental Petroleum Corp.	Energy	10
SLB	Schlumberger	Energy	10
XOM	Exxon Mobil Corp.	Energy	10
DD	DuPont	Materials	15
DOW	Dow Chemical	Materials	15
BA	Boeing Co.	Industrials	20
CAT	Caterpillar Inc.	Industrials	20
EMR	Emerson Electric Co.	Industrials	20
FDX	FedEx	Industrials	20
GD	General Dynamics	Industrials	20
GE	General Electric Co.	Industrials	20
HON	Honeywell	Industrials	20
LMT	Lockheed-Martin	Industrials	20
MMM	3M Company	Industrials	20
NSC	Norfolk Southern Corp.	Industrials	20
RTN	Raytheon Company	Industrials	20
UNP	Union Pacific Corp.	Industrials	20

UTX	United Technologies Corp.	Industrials	20
CMCSA	Comcast Corp.	Consumer Discretionary	25
DIS	The Walt Disney Company	Consumer Discretionary	25
F	Ford Motor	Consumer Discretionary	25
HD	Home Depot	Consumer Discretionary	25
LOW	Lowe's	Consumer Discretionary	25
MCD	McDonald's Corp.	Consumer Discretionary	25
NKE	Nike	Consumer Discretionary	25
SBUX	Starbucks Corp.	Consumer Discretionary	25
TGT	Target Corp.	Consumer Discretionary	25
TWX	Time Warner Inc.	Consumer Discretionary	25
CL	Colgate-Palmolive Co.	Consumer Staples	30
COST	Costco	Consumer Staples	30
CVS	CVS Caremark	Consumer Staples	30
KO	The Coca-Cola Company	Consumer Staples	30
MO	Altria Group	Consumer Staples	30
PEP	Pepsico Inc.	Consumer Staples	30
PG	Procter & Gamble Co.	Consumer Staples	30
WBA	Walgreens Boots Alliance	Consumer Staples	30
WMT	Wal-Mart	Consumer Staples	30
ABT	Abbott Laboratories	Health Care	35
AMGN	Amgen Inc.	Health Care	35
BAX	Baxter International Inc.	Health Care	35
BIIB	Biogen Idec	Health Care	35
BMJ	Bristol-Myers Squibb	Health Care	35
GILD	Gilead Sciences	Health Care	35
JNJ	Johnson & Johnson Inc.	Health Care	35

LLY	Eli Lilly and Company	Health Care	35
MDT	Medtronic Inc.	Health Care	35
MRK	Merck & Co.	Health Care	35
PFE	Pfizer Inc.	Health Care	35
UNH	UnitedHealth Group Inc.	Health Care	35
AIG	American International Group Inc.	Financials	40
ALL	Allstate Corp.	Financials	40
AXP	American Express Inc.	Financials	40
BAC	Bank of America Corp.	Financials	40
BK	Bank of New York	Financials	40
C	Citigroup Inc.	Financials	40
COF	Capital One Financial Corp.	Financials	40
JPM	JP Morgan Chase & Co.	Financials	40
MS	Morgan Stanley	Financials	40
SPG	Simon Property Group Inc.	Financials	40
USB	US Bancorp	Financials	40
WFC	Wells Fargo	Financials	40
AAPL	Apple Inc.	Information Technology	45
CSCO	Cisco Systems	Information Technology	45
EMC	EMC Corp.	Information Technology	45
IBM	International Business Machines	Information Technology	45
INTC	Intel Corp.	Information Technology	45
MSFT	Microsoft	Information Technology	45
ORCL	Oracle Corp.	Information Technology	45
QCOM	Qualcomm Inc.	Information Technology	45
TXN	Texas Instruments	Information Technology	45
T	AT&T Inc.	Telecommunications Services	50

VZ	Verizon Communications Inc.	Telecommunications Services	50
EXC	Exelon	Utilities	55
SO	Southern Company	Utilities	55

---

Table 6: Full data set

## C Descriptive Statistics of Historical Data

### Sector 10

Stock	Mean	St.Dev	Skewness	Kurtosis
APA	0.011537	0.100267	0.280947	0.760273
APC	0.012085	0.101281	0.466194	1.93809
COP	0.010643	0.071961	-0.05553	1.251128
CVX	0.010086	0.0586	0.153503	0.806617
DVN	0.010759	0.097821	0.10065	1.082427
HAL	0.013205	0.112151	-0.32884	1.382529
OXY	0.01384	0.080393	0.349282	1.099848
SLB	0.012642	0.092495	-0.13044	1.233693
XOM	0.009766	0.049339	0.349948	1.479707

---

### Sector 15

Stock	Mean	St.Dev	Skewness	Kurtosis
DD	0.009388	0.075868	0.296421	1.275727
DOW	0.011763	0.106412	2.365318	19.95742

---

**Sector 20**

Stock	Mean	St.Dev	Skewness	Kurtosis
BA	0.012435	0.080764	-0.62964	1.301433
CAT	0.013029	0.093177	0.252294	2.670581
EMR	0.008808	0.065555	-0.08548	0.480123
FDX	0.012659	0.081811	0.466013	1.635417
GD	0.014238	0.068401	-0.25389	1.993718
GE	0.010362	0.075521	-0.21698	1.403417
HON	0.012723	0.087006	-0.21906	6.623444
LMT	0.013811	0.071289	-0.98771	4.287257
MMM	0.01103	0.06064	0.225859	1.377233
NSC	0.010869	0.081867	0.267064	0.620407
RTN	0.010583	0.07996	-0.47971	4.586874
UNP	0.012855	0.068637	-0.01567	0.830352
UTX	0.013932	0.068338	-0.58433	3.092721

**Sector 25**

Stock	Mean	St.Dev	Skewness	Kurtosis
CMCSA	0.013679	0.080699	-0.09995	0.263036
DIS	0.011032	0.07378	-0.25131	1.278885
F	0.012278	0.137978	3.04211	28.61992
HD	0.014443	0.077528	0.087253	0.668349
LOW	0.015936	0.089288	0.237749	0.208749
MCD	0.011475	0.060913	-0.39298	1.57291
NKE	0.017885	0.085516	-0.20359	3.807669

SBUX	0.023437	0.104151	-0.18577	1.967423
TGT	0.014447	0.078877	0.192496	0.650608
TWX	0.022785	0.133481	1.318255	5.810858

---

**Sector 30**

Stock	Mean	St.Dev	Skewness	Kurtosis
CL	0.012162	0.063406	0.565845	4.617323
COST	0.016273	0.072085	-0.68088	4.405728
CVS	0.014408	0.07461	-0.33915	1.766361
KO	0.008461	0.06076	-0.29419	1.495421
MO	0.017196	0.072493	-0.21977	2.765067
PEP	0.010451	0.0565	-0.36449	3.978095
PG	0.009936	0.056466	-0.91	7.43439
WBA	0.014574	0.077105	0.296946	0.524327
WMT	0.009982	0.064803	0.234029	1.191134

---

**Sector 35**

Stock	Mean	St.Dev	Skewness	Kurtosis
ABT	0.010797	0.056754	-0.48062	0.565874
AMGN	0.015989	0.086528	0.851517	2.459323
BAX	0.010978	0.070055	-0.54315	3.221924
BIIB	0.038001	0.157041	0.862778	3.515306
BMJ	0.011579	0.068866	-0.32027	1.30521
GILD	0.029605	0.125628	0.933879	3.487074
JNJ	0.011129	0.053357	0.056419	0.869805

LLY	0.011795	0.074798	0.100057	2.302815
MDT	0.012614	0.067606	-0.05722	1.533405
MRK	0.00962	0.075546	-0.10188	0.666631
PFE	0.010597	0.063625	-0.09411	0.020745
UNH	0.016105	0.085046	-1.01332	3.244448

---

**Sector 40**

Stock	Mean	St.Dev	Skewness	Kurtosis
AIG	0.012391	0.214508	6.515358	73.88875
ALL	0.011989	0.080738	-0.86063	5.182934
AXP	0.013126	0.093359	2.739294	28.95565
BAC	0.010432	0.113328	0.333408	9.134332
BK	0.011623	0.080714	0.247555	1.189955
C	0.00879	0.121562	0.129994	8.026639
COF	0.0183	0.118748	0.073638	3.511781
JPM	0.013361	0.093858	-0.19985	1.278393
MS	0.013166	0.111323	-0.01598	2.425499
SPG	0.015641	0.072354	0.20246	9.266868
USB	0.011719	0.077209	-0.90354	5.708111
WFC	0.014313	0.081502	0.117893	5.471074

---

**Sector 45**

Stock	Mean	St.Dev	Skewness	Kurtosis
AAPL	0.026328	0.129514	-0.3117	1.988668
CSCO	0.017094	0.107572	-0.14407	1.257588

EMC	0.017846	0.125136	-0.07486	1.396756
IBM	0.012237	0.078643	0.491961	2.286611
INTC	0.015574	0.10727	-0.31768	1.428539
MSFT	0.016669	0.095078	0.473716	2.140453
ORCL	0.019108	0.125136	0.954885	4.582381
QCOM	0.024179	0.148131	1.875931	9.102131
TXN	0.017676	0.115362	0.38084	1.79874

---

**Sector 50**

Stock	Mean	St.Dev	Skewness	Kurtosis
T	0.007774	0.067997	0.186248	1.484053
VZ	0.008348	0.06797	0.821935	4.251545

---

**Sector 55**

Stock	Mean	St.Dev	Skewness	Kurtosis
EXC	0.008614	0.067595	0.110001	2.3338
SO	0.010475	0.049348	0.317289	2.40358

Table 7: Descriptive statistics of Historical Data

## D Sector Comparison

### D.1 Variance Explained by PC

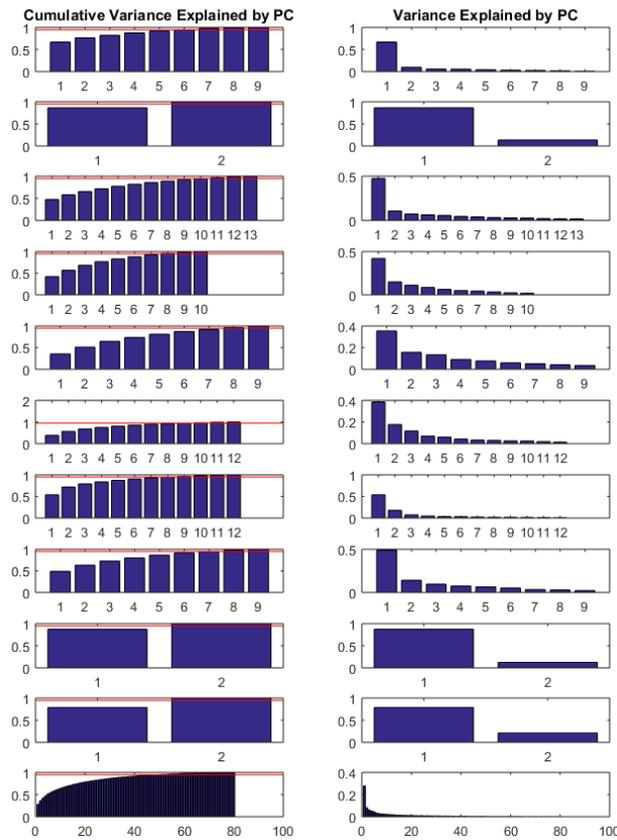


Figure 21: Cumulative variance and variance explained by principal component for all sectors and the full data set, respectively

## D.2 Correlation Matrices

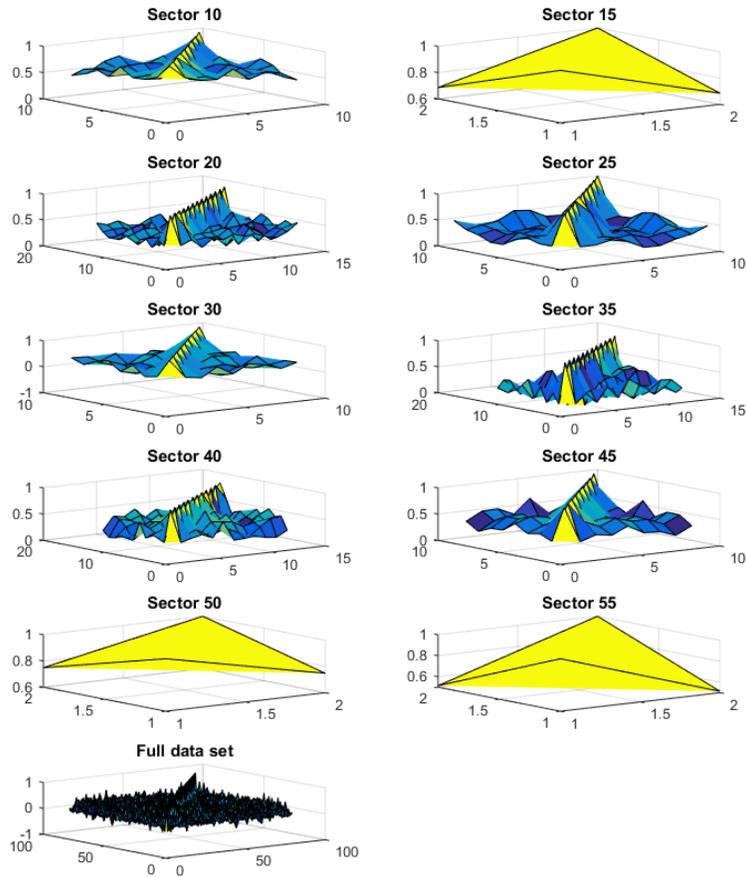


Figure 22: Correlation matrices for all sectors and for the full data set

## E Explanation of Numerical Instability

**Theorem 1** *If  $X_1, X_2, \dots, X_n$  are mutually independent normally distributed random variables with means  $\mu_1, \mu_2, \dots, \mu_n$  and variances  $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$ , then the linear combination*

$$Y = \sum_{i=1}^n c_i X_i \quad (44)$$

*follows the normal distribution:*

$$N\left(\sum_{i=1}^n c_i \mu_i, \sum_{i=1}^n c_i^2 \sigma_i^2\right) \quad (45)$$

This theorem applies to the data points  $Q_{r_i}(q|PC_1, \dots, PC_m)$  as the PCs are mutually independent normally distributed when simulated using GBM. Thus, the variance of each data point becomes larger when  $n$  grows.

# F January 2016 CI Analysis

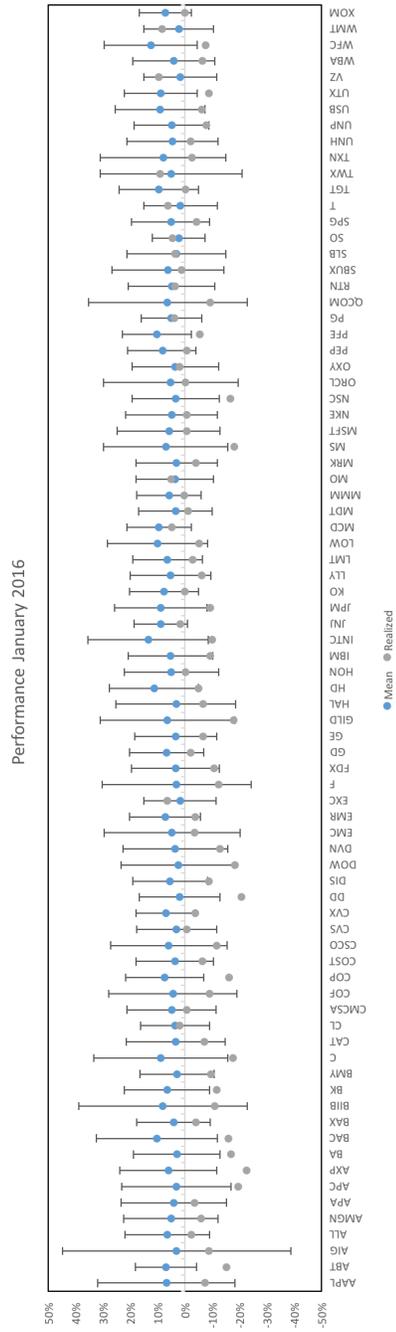


Figure 23: Realized and mean returns

# G Simulation Plots

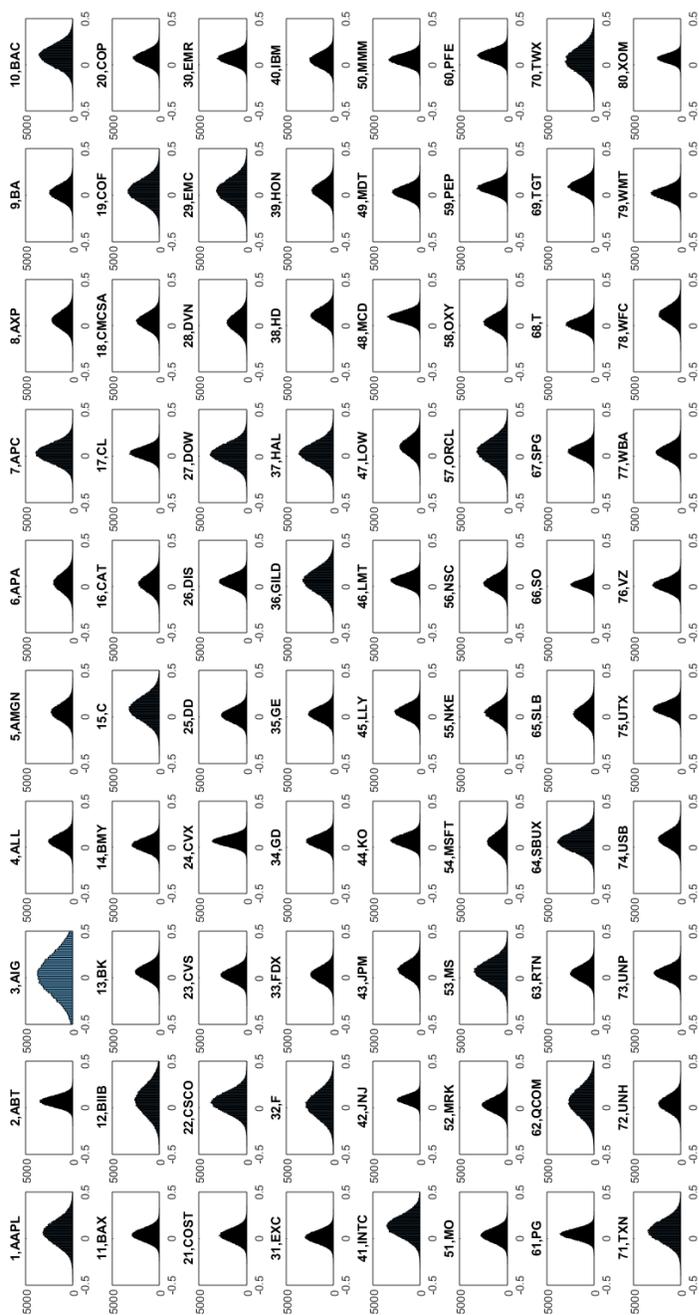
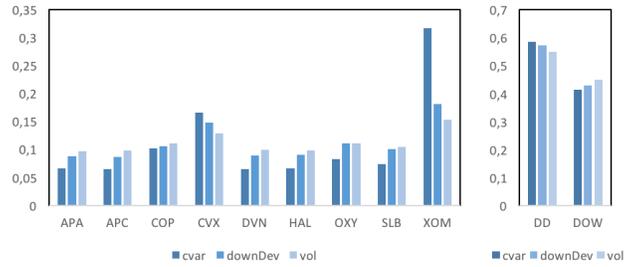


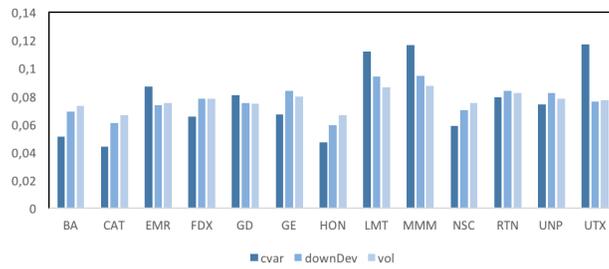
Figure 24: Return distributions for all stocks with 100,000 simulations

# H Risk Parity Portfolio Weights

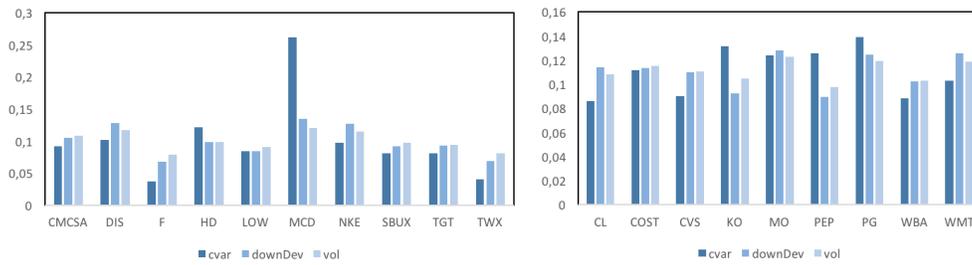


(a) Sector 10

(b) Sector 15

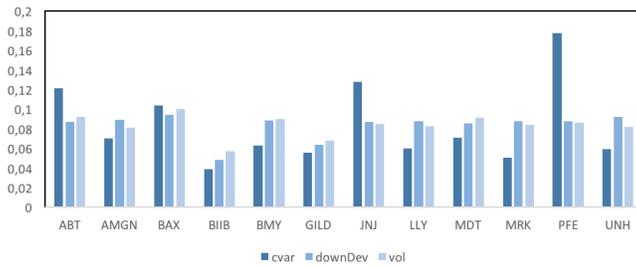


(c) Sector 20

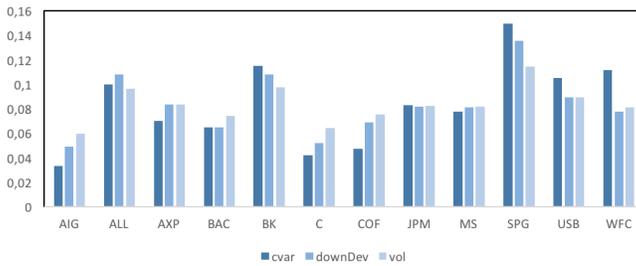


(d) Sector 25

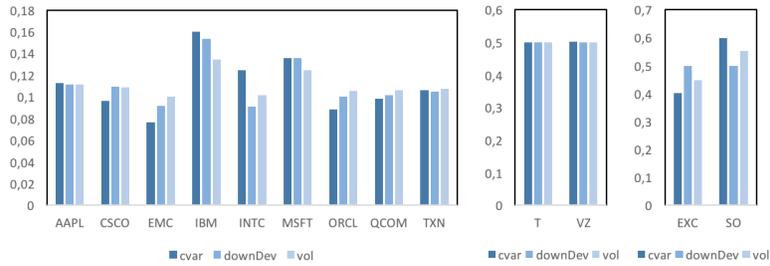
(e) Sector 30



(f) Sector 35



(g) Sector 40



(h) Sector 45

(i) Sector 50 (j) Sector 55

Figure 24: Risk parity weight allocations for all sectors using three different risk measures