CHARLES UNIVERSITY IN PRAGUE

FACULTY OF SOCIAL SCIENCES
Institute of Economic Studies

Inés Horváthová

Robust portfolio selection

Bachelor thesis

Prague 2014

Author: **Inés Horváthová**

Supervisor: **RNDr. Michal Červinka, Ph.D.**

Year of defense: **2014**

Bibliografický záznam

HORVÁTHOVÁ, Inés. *Robust portfolio selection*. Praha 2014. 59 s. Bakalářská práce (Bc.) Univerzita Karlova, Fakulta sociálních věd, Institut ekonomických studií a žurnalistiky. Vedoucí bakalářské práce: RNDr. Michal Červinka, Ph.D.

Character count: 60 016

Anotace (abstrakt)

V předložené práci budeme studovat optimalizaci portfolií pomocí „mean-risk" modelů. Nejdříve si obecně definujeme rizikové míry a dále uvedeme tři běžně užívané: rozptyl, Value-at-risk (VaR - „hodnota v riziku") a Conditional value-at-risk (CVaR - „podmíněná hodnota v riziku"). Pro každou z těchto rizikových měr formulujeme příslušné „mean-risk" modely. Dále ke každé z nich uvedeme robustní verzi. Nejvíce se budeme věnovat robustním verzím modelů s rozptylem jako mírou rizika, které následně aplikujeme na historická data, využitím volně dostupného statistického softwaru R. Nakonec porovnáme získané výsledky s klasickým nerobustním modelem mean-variance.

Klíčová slova:	výběr portfolií, VaR, CVaR, mean-variance, robustní
E-mail autora:	ines.horvathova@gmail.com
E-mail vedoucího práce:	cervinka@utia.cas.cz

Abstract

In this thesis, we take the mean-risk approach to portfolio optimization. We will first define risk measures in general and then introduce three commonly used ones: variance, Value-at-risk (VaR) and Conditional-value-at-risk ($CVaR$). For each of these risk measures we

formulate the corresponding mean-risk models. We then present their robust counterparts. We focus mainly on the robust mean-variance models, which we also apply to historical data using free statistical software R. Finally, we compare the results with the classical non-robust mean-variance model.

Declaration of Authorship

I hereby proclaim that I wrote my bachelor thesis on my own under the leadership of my supervisor and that the references include all resources and literature I have used.

I further declare that the thesis has not been used previously for obtaining any university degree.

I grant a permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, July 17, 2014

_____

Signature

Acknowledgment

I would like to express my gratitude to the supervisor of my bachelor thesis RNDr. Michal Červinka, Ph.D. for his time, patience, advice and guidance.

Bachelor Thesis Proposal

In the thesis, the author will first introduce risk measures and then analyze three commonly used risk measures: variance, Value at Risk (VaR) and Conditional Value at Risk ($CVaR$). For each of these three risk measures she will formulate corresponding portfolio selection models: classical mean-variance model, mean-VaR model and mean-$CVaR$ model. Next, the author will specify robust counterparts of classical mean-variance model. In the last part of the thesis, the author will illustrate robust portfolio selection by applying classical mean-variance model and its robust counterparts on historical data and compare the results. The theoretical part of the thesis will be based primarily on [1] and references therein.

Preliminary structure of the thesis

1. Introduction
2. Risk measures and preliminaries
3. Portfolio selection models

   3.1. Mean-Variance model
   3.2. Mean-VaR model
   3.3. Mean-$CVaR$ model
   3.4. Robust portfolio selection models

4. Numerical example
5. Conclusion

Literature

[1] F. J. Fabozzi, D. Huang, G. Zhou: Robust portfolios: contributions from operations research and finance, Annals of Operations Research 176(1), 191 - 220, 2010.

[2] R. Tütüncü, M. Koenig: Robust asset allocation. Annals of Operations Research 132, 157 - 187, 2004.

Contents

Chapter 1

Introduction

Investment is, in essence, present sacrifice for future benefit. But the present is relatively well known, whereas the future is always an enigma. Investment is also, therefore, certain sacrifice for uncertain benefit.

<div align="right">(Jack Hirshleifer, 1965)</div>

An investor decides how to invest, so that the resulting investment strategy is in some sense optimal. In general, there are two types of decision making frameworks used in order to find this optimal strategy: the utility maximization and the return-risk analysis. Utility maximization was first introduced by Neumann and Morgenstern in 1944 [18]. The optimal strategy is chosen so that the investor's utility is maximized. Despite being very profound in theory, this approach describes the risk in an indirect way, and consequently is rarely used in practice. Moreover, as is written in [6] for similar levels of risk aversion, different utility functions result in similar asset allocations. Hence, misspecification of utility function does not change the optimal portfolio much.

For these reasons, in this thesis we will take the second approach to portfolio optimization, the return-risk analysis, which is broadly used both in theory and practice. In this approach, an investor faces a trade-off between expected return and associated risk. The risk is clearly quantified by a risk measure that maps the loss to a real number. The return-risk approach was pioneered by Markowitz in 1952 [15], who adopted variance as the measure of risk in his mean-

variance analysis. Although it is simple to use and interpret, variance measures equally both positive and negative fluctuations. In finance, however, attention is mostly given to losses, and hence, downside risk measures are preferred. Already Markowitz himself was aware of this shortcoming of variance and in [16] in 1959 he proposes a measure of downside risk, called semi-variance.

Since then, many other risk measures were introduced. In this thesis, we will first introduce risk measures in general and then define already mentioned variance and two commonly used risk measures: Value-at-risk (VaR) and Conditional-value-at-risk ($CVaR$) in Chapter 3. VaR is a quantile risk measure, widely popular in banking and insurance sector, which measures the maximum loss that can be expected during a specific time horizon with a probability of α (α close to 1). Despite of its popularity, VaR lacks some important mathematical properties and does not provide information about the losses that are to occur with small probability but with disastrous consequences [2]. So new measures based on VaR were introduced. One of them is $CVaR$ [22], which is defined as the mean value of losses that exceed the value of VaR. $CVaR$ exhibits many favourable mathematical properties and its popularity is rising.

For each of these risk measures, we will then formulate corresponding mean-risk optimization problems in Chapter 4. These models are strongly dependent on the parameters of the underlying distribution, which are never known for certain and hence have to be estimated [8]. Investors usually rely on limited data to estimate the input parameters and so estimation errors are unavoidable. A small change in the input parameters should result in no or only in a small change in the optimal asset allocation [30]. This is, however, not always true and changes in input parameters can result in significant changes in the solutions to different portfolio selection models. Hence, an estimation risk should also be incorporated into the problem of searching for an optimal portfolio. Robust portfolio selection deals with eliminating the impacts of

2

this estimation risk. There exist different robust approaches to the portfolio optimization [8].

Chapter 5 is devoted to introducing robust counterparts to the previously formulated mean-risk models. The focus is mainly put on the mean-variance model, for which we formulate three robust counterparts: uncertainty in the estimate of mean only, in the estimate of variance only and then in both the estimates of mean and variance. In reference to [6], the estimates of covariances are the least important in terms of their influence on the optimal portfolio and so covariances are assumed to be true. The classical mean-variance problem along with its robust counterparts are then calculated for historical data in Chapter 6. We then show the resulting asset allocations and compare the results.

Chapter 2

Preliminaries and basic definitions

A portfolio is defined as a collection of (financial) assets. A rational investor chooses their portfolio, so that the expected return is maximized and the risk is minimized. This concept of portfolio optimization was first introduced into modern finance by Markowitz in 1952 [15]. Measures of risk have a crucial role in coping with the losses that might be incurred in finance [23].

Let us assume there are N risky financial assets and one risk-free asset an investor can choose from. We denote $\mathcal{X} \subset \mathbb{R}^{N+1}$ a set of possible decisions how disposable wealth can be allocated. Every decision is given by a vector $\mathbf{x} \in \mathcal{X}$. These decisions in fact represent portfolios and each component x_i represents the fraction of disposable wealth to be invested into the i-th asset, $i = 0, 1, \ldots, N$. We do not allow *short-sales*, i.e. the situation when an investor can sell a security that they do not own. Hence, we assume $x_i \geq 0$ for every $i = 0, 1, \ldots, N$. The feasible set of portfolios $\mathcal{X} \subset \mathbb{R}^{N+1}$ is then of the form:

$$\mathcal{X} = \left\{ \mathbf{x} \in \mathbb{R}^{N+1} | \sum_{i=0}^{N} x_i = 1, \, x_i \geq 0 \right\},$$

i.e. no short-sales are allowed and all the money available for investment is allocated.

Furthermore, we consider only one-period decision horizon. For $T > 0$ we define a random return for period $[0, T]$ as $\boldsymbol{\rho} = (\rho_0, \rho_1, ..., \rho_N)^T$,

where a one-unit investment into the i-th risky asset yields return of ρ_i, $i = 1, ..., N$ in $[0, T]$ and the return of risk-free asset is ρ_0. The distribution of random vector $\boldsymbol{\rho}$ is independent of $\mathbf{x}$ and is given by the expected return $\mathbf{r} = \mathbb{E}\boldsymbol{\rho} = (r_0, r_1, ..r_N)^T$ and the covariance matrix $\mathbf{V} = (\sigma_{ij})$, where $\sigma_{ij} = cov(\sigma_i, \sigma_j)$, $i,j = 0, 1, ...N$, [7] and [30].

Definition 2.1. *Let $\boldsymbol{\rho}$ be a random vector of returns and $\boldsymbol{x} \in \mathcal{X} \subset \mathbb{R}^{N+1}$. We define a return of a portfolio as a random variable*

$$Y = \boldsymbol{x}^T \boldsymbol{\rho}$$

and the expected return of a portfolio as

$$R = \boldsymbol{x}^T \boldsymbol{r}.$$

Let us denote $\mathcal{Y}$ a set of random returns, that is $\mathcal{Y} = \{Y = \mathbf{x}^T \boldsymbol{\rho} | x \in \mathcal{X}\}$.

Intuitively, a negative return is a loss. Let us now define a loss function.

Definition 2.2. *Loss function is defined as a random variable $L = -\boldsymbol{x}^T \boldsymbol{\rho}$, which is a function of a decision vector $\boldsymbol{x} \in \mathcal{X} \subset \mathbb{R}^{N+1}$ and a random variable $\boldsymbol{\rho} = (\rho_0, \rho_1, ..., \rho_N)^T$.*

Distribution of the loss function depends on the choice of $\mathbf{x}$. Let us denote $\mathcal{L}$ a set of random losses, where $\mathcal{L} = \{L = -\mathbf{x}^T \boldsymbol{\rho} | x \in \mathcal{X}\}$.

Now we will define a measure of risk.

Definition 2.3. *[20] Let $Z \in \mathcal{Z}$ be a random variable, either random loss $L \in \mathcal{L}$ or random return $Y \in \mathcal{Y}$. Then risk measure is a probability functional $\mathcal{R} : \mathcal{Z} \to \mathbb{R}$.*

Higher values of this measure represent higher risk and hence a rational investor prefers lower values. In the next chapter we will describe properties of risk measures that are often required and then introduce three commonly used risk measures: variance, Value-at-Risk (VaR) and Conditional Value-at-Risk ($CVaR$).

Chapter 3

Risk measures

General risk measure is defined as a functional that assigns a real number to a random variable, which can be either loss or return. Following [2] we now state properties of risk measures that are required most often along with their economic interpretations.

(S) Subadditivity, i.e. $\rho(X + Y) \leq \rho(X) + \rho(Y)$ for any X, Y

This property has a clear economic interpretation. Let us assume that X and Y are two different portfolios and $\rho(X)$, $\rho(Y)$ represent their risks. Fusing these two portfolios into a single portfolio $Z = X + Y$ and applying property (S), it has to hold that the risk of portfolio Z is no greater than the sum of risks of the two original portfolios. In other words, this property guarantees that diversifying portfolio does not create extra risk.

(PH) Positive homogeneity, i.e. $\rho(\lambda X) = \lambda \rho(X)$ for any $\lambda \geq 0, X$

This property simply says that the risk of a portfolio that is λ times bigger than our original portfolio equals λ times the risk of the original portfolio. This property is closely related to liquidity, as the higher our position is, the less liquid it is and hence it is associated with higher risk.

(M) Monotonicity, i.e. $X \leq Y \, a.e. \Rightarrow \rho(X) \geq \rho(Y)$

Let us assume that random variables X and Y represent the values of portfolios A and B, respectively. Then relation $X \leq Y \, a.e.$ means that

portfolio B has higher value than portfolio A with probability equal to 1. If property (M) holds, then the risk of portfolio B (measured by $\rho(Y)$) is less than the risk of portfolio A (measured by $\rho(X)$).

(TI) Translation invariance, i.e. $\rho(X + \alpha) = \rho(X) - \alpha$ for any X and $\alpha \in \mathbb{R}$

This property suggests that if we add a risk-free constant value of $\rho(X)$ to our financial position X, then our resulting investment is risk-free, as $\rho(X + \rho(X)) = \rho(X) - \rho(X) = 0$. Under property (TI) the risk measure $\rho(X)$ actually tells us how much constant and risk-free money we have to add to our risky position, so that it becomes risk-free. In other words, this property ensures that the resulting risk measured by a risk measure is expressed in the same units as risk-free alternative to risky investment: money.

(C) Convexity, i.e. $\rho(\lambda X + (1 - \lambda)Y) \leq \lambda \rho(X) + (1 - \lambda)\rho(Y)$ for any X, Y and $0 \leq \lambda \leq 1$

Its economic interpretation is similar to the one of property (S). Moreover, assuming property (PH) holds, properties (S) and (C) can be used interchangeably. This fact can be shown as follows:
If properties (S) and (PH) hold, then $\rho(\lambda X + (1 - \lambda)Y) \leq \rho(\lambda X) + \rho((1 - \lambda)Y) = \lambda \rho(X) + (1 - \lambda)\rho(Y)$ for any X, Y and $0 \leq \lambda \leq 1$ and so property (C) holds.
On the other hand, when properties (C) and (PH) hold, then $\rho(X + Y) = \rho(\lambda/\lambda\, X) + \rho((1 - \lambda)/(1 - \lambda)\, Y) \leq \lambda \rho(1/\lambda\, X) + (1 - \lambda)\rho(1/(1 - \lambda)\, Y) = \lambda/\lambda\, \rho(X) + (1 - \lambda)/(1 - \lambda)\, \rho(Y) = \rho(X) + \rho(Y)$ and hence property (S) holds.

(SI) Shift invariance, i.e. $\rho(X + \alpha) = \rho(X)$ for any X and $\alpha \in \mathbb{R}$

In this case we are looking at the risk as a degree of uncertainty associated with unfavourable events and situations. Favourable situations do not cause any danger to us and hence we are not interested in them.

(N) Nonnegativity, i.e. $\rho(X) > 0$ for nonconstant X and $\rho(X) = 0$ for constant X

In this case we also look at the risk as a degree of uncertainty associated with unfavourable events. When X is constant, there is no risk and hence $\rho(X) = 0$. If X is not a constant, then there is always some possibility of a future unfavourable situation or result and hence $\rho(X) > 0$.

(EB) Expectation bounded, i.e. $\rho(X) > \mathbb{E}[-X]$ for nonconstant X and $\rho(X) = \mathbb{E}[-X]$ for constant X

This property does not have a clear economic interpretation. In some sense the first four properties are the most important and the most desirable, as they represent all crucial intuitive and practical properties one may require a risk measure to possess. So in [4] Artzner et al. use them to define a special type of risk measures, which are called coherent.

Definition 3.1 (Coherent risk measure). *A risk measure satisfying the four properties of translation invariance, subadditivity, positive homogeneity, and monotonicity is called coherent.*

There are many different risk measures that are coherent. However, there also exist some risk measures that are used in practice and yet lack the coherence property. One of these non-coherent risk measures is a broadly used VaR - "Value at Risk". VaR satisfies properties of (PH), (M) and (TI), but lacks property (S). VaR is not subadditive and hence not coherent. This fact is a great disadvantage of this risk measure. More on this in the section devoted to VaR.

Risk measures that are not coherent can still belong to a class of convex risk measures.

Definition 3.2 (Convex risk measure). *A risk measure satisfying the three properties of monotonicity, translation invariance and convexity is called a convex risk measure.*

As shown above, if properties (S) and (PH) hold, also property (C) holds and hence every coherent risk measure is also convex. The other implication does not hold, as we would need to have property (PH) along with property (C) in order to get property (S). So the class of convex risk measures include all coherent risk measures and some other, hence convex risk measures are sometimes referred to as weakly coherent risk measures.

There is another possibility how to quantify risk, specifically there exists a class of deviation risk measures, which we will define according to [2].

Definition 3.3 (Deviation risk measure). *A risk measure $\mathcal{D}(X)$ that satisfies the four properties of subadditivity, positive homogeneity, shift invariance and nonnegativity is called a deviation risk measure.*

A deviation risk measure $\mathcal{D}(X)$ depends only on fluctuations around the mean value, which can be seen from the property (SI), as $\mathcal{D}(X - \mathbb{E}[X]) = \mathcal{D}(X)$. An example of such a deviation risk measure is standard deviation.

As is shown in [20], no risk measure can be both deviation and coherent. This results from the fact that the property (TI), required for coherency, and the property (SI), required for a risk measure to be a deviation risk measure, contradict each other. However, these two classes of risk measures are related in such a way that one can easily construct one type of risk measure from the other type by appropriately including the mean value to the construction. But this relation does not hold for all coherent risk measures, it holds only for so called strictly expectation bounded risk measures.

Definition 3.4 (SEBRM). *A risk measure $\mathcal{R}(X)$ that satisfies the four properties of subadditivity, positive homogeneity, translation invariance and expectation bounded is called a strictly expectation bounded risk measure.*

These SEBRM risk measures differ from coherent risk measures as they satisfy an additional property of (EB) and do not strictly satisfy the property of (M). An example of such a risk measure is $\mathcal{R}(X) = \lambda\sigma[X] - \mathbb{E}[X]$ for any $\lambda > 0$.

Now we state how deviation risk measures and SEBRM are connected in reference to [2].

Proposition 3.1. *If $\mathcal{R}(X)$ is a SEBRM risk measure, then $\mathcal{D}(X) = \mathcal{R}(X - \mathbb{E}[X])$ is a deviation risk measure and if $\mathcal{D}(X)$ is a deviation risk measure, then $\mathcal{R}(X) = \mathcal{D}(X) - \mathbb{E}[X]$ is a SEBRM risk measure.*

In the next section we will introduce and describe three commonly used measures of risk, namely variance, Value-at-Risk (VaR) and Conditional Value-at-Risk ($CVaR$).

3.1   Variance and standard deviation

We define variance and standard deviation as measures of portfolio risk according to Markowitz [15]. He focuses at the correlation of returns between individual assets; $cov(\rho_i x_i, \rho_j x_j) = x_i cov(\rho_i, \rho_j) x_j$, where covariances $cov(\rho_i, \rho_j)$ for $i, j = 0, 1, \ldots, N$ are elements of the variance matrix $\mathbf{V}$. He shows that keeping return of a portfolio unchanged, the lower the correlation of individual asset returns is, the lower is the risk of the entire portfolio measured using variance or standard deviation.

Definition 3.5. *The risk of a portfolio with weights $\boldsymbol{x}$ measured by variance is defined as the variance of the random portfolio return*

$$\sigma^2(\boldsymbol{x}) = \sum_{i,j}[cov(\rho_i, \rho_j)]x_i x_j = \boldsymbol{x}^T \boldsymbol{V}\boldsymbol{x}.$$

Variance is a symmetric risk measure, $\sigma^2(\mathbf{x}) = \sigma^2(-\mathbf{x})$. It does not belong to the class of deviation risk measures, as it fails to satisfy the

property (PH), $\sigma^2(\lambda\mathbf{x}) = \lambda^2\sigma^2(\mathbf{x}) \neq \lambda\sigma^2(\mathbf{x})$. Sometimes the standard deviation is taken instead of the variance as a risk measure.

Definition 3.6. *The risk of a portfolio with weights $\boldsymbol{x}$ measured by standard deviation is defined as the standard deviation of the random portfolio return*

$$\sigma(\boldsymbol{x}) = \sqrt{\sum_{i,j}[cov(\rho_i, \rho_j)]x_i x_j} = \sqrt{\boldsymbol{x}^T \boldsymbol{V} \boldsymbol{x}}.$$

Standard deviation is also a symmetric risk measure, $\sigma(\mathbf{x}) = \sigma(-\mathbf{x})$, which satisfies all four properties (S), (PH), (SI), (N) and hence it is a deviation risk measure.

Both standard deviation and variance consider both positive and negative fluctuations of returns and hence they provide a measure of volatility. In finance, however, extremely high returns are not considered to be a danger. On the other hand, these extremely positive events are often welcome. This creates a limitation for variance and standard deviation, which can be eliminated by simply incorporating only unfavourable movements into the calculation of variance. Already Markowitz himself was aware of this shortcoming of variance and in [16] proposes a measure of downside risk, called semi-variance. There also exist one-sided versions of standard deviation, which still belong to the class of deviation risk measures, but are not symmetric.

3.2   Value-at-Risk

This measure of risk was developed in the investment bank J. P. Morgan, where it was used to provide the management with a daily brief and clear information about the risk and losses that are to be reasonably expected the next day. Thanks to its simplicity and the fact that J. P. Morgan provided elaborate presentations and technical instructions to banks and financial institutions, Value-at-Risk (VaR)

as a measure of risk started to be used broadly. This popular risk measure became a standard in the banking regulation Basel I, Basel II and slightly improved and modified also in Basel III. VaR is also a part of Solvency II, regulation for insurance companies.

VaR gives an answer to the following question:
What is the maximum loss of a portfolio that we can expect during a specific time horizon with a probability of α?

Or alternatively:
What is the minimum loss of a portfolio that we can expect during a specific time horizon in the worst-case scenarios that occur with a probability of $(1 - \alpha)$?

From these questions we can see that VaR depends on two parameters:

- Holding period, which specifies in what time period potential losses will be considered. One day, one week, ten days, one month and one year are used most often.

- Confidence level, which states how probable/ improbable events will be taken into account. It can be given in percentage or using $\alpha \in (0, 1)$, where $\alpha \cdot 100 =$ confidence level in %. Often used are confidence levels of 99.9%, 99% or 95%. In practice, for example, Bankers Trust uses $\alpha = 0.99$, J. P. Morgan $\alpha = 0.95$ and Citibank[1] $\alpha = 0.954$.

In order to define VaR properly for a general distribution function we will use *lower* and *upper* α-quantiles and so we will first state their formulas:

$$
\begin{aligned}
lower\ \alpha\text{-quantile} \quad &\sup\{x \mid F_X(x) < \alpha\} = q_\alpha(X) \\
upper\ \alpha\text{-quantile} \quad &\inf\{x \mid F_X(x) > \alpha\} = q^\alpha(X),
\end{aligned}
\tag{3.1}
$$

where $F_X(x) = P[X \leq x]$ is a distribution function of a random variable X.

---

[1] 95.4% is associated with ± 2 standard deviations of normal distribution.

Alternatively, we can write

$$\inf\{x \mid F_X(x) \geq \alpha\} = \sup\{x \mid F_X(x) < \alpha\} \equiv q_\alpha(X)$$

and

$$\sup\{x \mid F_X(x) \leq \alpha\} = \inf\{x \mid F_X(x) > \alpha\} \equiv q^\alpha(X)$$

From this we can see a relation between *lower* and *upper* α-quantiles:

$$q_\alpha(X) = \sup\{x \mid F_X(x) < \alpha\} \leq \sup\{x \mid F_X(x) \leq \alpha\} = q^\alpha(X),$$

where equality holds only for continuous and strictly increasing distribution functions. Now we can define VaR_α as follows:

Definition 3.7 (VaR). *Let $X \in \mathcal{Y}$ be a random return of a portfolio and $\alpha \in (0,1)$. Value-at-Risk VaR_α of the portfolio associated with a decision $\boldsymbol{x} \in \mathcal{X}$ is defined as*

$$VaR_\alpha(\boldsymbol{x}) = -q_\alpha(X).$$

In the definition of VaR we use a minus sign, so that we get loss as a positive number. We can see that VaR_α is just the minus *lower* α-quantile of the distribution function for random return.

In order to deal with general or discrete distribution functions an upper VaR_α^+ was introduced.

Definition 3.8 (VaR^+). *The upper VaR_α^+ of the portfolio associated with a decision $\boldsymbol{x} \in \mathcal{X}$ is defined as*

$$VaR_\alpha^+(\boldsymbol{x}) = -q^\alpha(X).$$

From the definitions of VaR and upper VaR^+ and the relation between *lower* and *upper* α-quantile we can see that relation $VaR_\alpha(X) \geq VaR_\alpha^+$ always holds.

As written above, VaR as a measure of risk became quickly popular and widely used. A lot of books and articles describing its positives

and various applications were published. According to [2], we will now present some strengths of VaR.

1. Simplicity. Method of VaR is very simple to use and easy to understand. This might be the reason for its wide use.

2. Unit of measurement. VaR represents "the money that is lost", i.e. its unit of measurement is currency (USD, Euro or any other), which makes this risk measure easy to grasp.

3. Stability. When it comes to estimation processes VaR behaves rather stable. Because it disregards the tails, it is not affected by very high tail losses, which are usually difficult to measure [25]. On the other hand, this ignorance of tail is also one of VaR's disadvantages and in this context will be discussed below.

4. Universality. VaR can be used to measure the risk of any portfolio: bond portfolios, stock portfolios, mixed portfolios, etc. VaR is mostly used to quantify the market risk, but it can be used measure any type of risk.

5. Generality. A portfolio may be exposed to many types of risk: interest rate risk, exchange rate risk, etc. VaR summarizes these risks and measures them together at one point in time. VaR is hence a general risk measure.

6. Down-side risk measure. VaR takes into account only negative events, as extremely positive returns are not considered a risk in finance, on the contrary, are often welcome. This gives VaR an advantage compared to variance.

7. Probability. VaR provides information about the probabilistic behaviour of the entire portfolio. Specifying VaR for all confidence levels completely defines the distribution. In this sense, VaR is superior to standard deviation [25].

Despite of its popularity, there are many drawbacks to VaR. These drawbacks were gradually uncovered and the criticism of VaR peaked in the period of financial crisis 2007-2009, when it was actually listed as one of the prime culprits of the crisis. We will now state some of its weaknesses. We follow [2], where an interested reader may find relevant academic examples.

1. Subadditivity. VaR lacks the subadditivity (S) property and hence, in general, it is not a coherent measure of risk [3]. This non-subadditivity suggests that risk can be eliminated by dividing a single portfolio into many smaller parts. It can actually be shown that if a company is free to diversify its portfolio into any number of portfolios, then under mild assumptions every portfolio can be divided into smaller sub-portfolios, each with VaR equal to 0 [10]. In practice, however, we expect that joining risky portfolios into one would eliminate part of the risk and hence that the risk of the joined portfolio would not be greater than the sum of risks of individual portfolios. Nevertheless, for some probability distributions, for example, for elliptical distribution[2], is VaR subadditive and hence coherent. For normal distributions VaR is proportional to the standard deviation [22], [4].

2. Convexity. VaR is generally not even convex (weakly coherent). This non-convexity of VaR causes difficulties mainly when calculating various optimization problems and hence minimization of this risk measure becomes hard to solve.

3. VaR and correlation. According to Markowitz, the higher the correlation between random variable is, the higher should be the value of a risk measure. VaR, however, does not satisfy this requirement.

---

[2]Among elliptical distributions belong, for example, normal distribution, t-distribution, logistic distribution, etc.

4. VaR and normal distribution. VaR implicitly assumes that the random variable is normally distributed, although in reality it is not the case. Extremely improbable events that lead to extreme losses have already happened a number of times.[3] In practice distributions tend to exhibit empirical discreteness or "fat tails".

5. Values beyond VaR. VaR_α chooses the least risky scenario out of all $\alpha \cdot 100\%$ worst-case scenarios. Under the worst-case scenarios we mean the largest losses in the holding period that occur with probability less than α. VaR systematically underestimates the potential risk and hence it is biased toward optimism instead of conservatism that should prevail in risk management. Moreover, it does not provide us with the value of the losses that might be suffered with a small probability, but with disastrous consequences. This ignorance of tail actually implies that $VaR_\alpha(\mathbf{x})$ may increase dramatically with a small increase in α.

6. Inconsistent results. Basic methods of calculating VaR can be divided into two main categories, parametric and non-parametric methods. Parametric methods are based on statistical parameters, for example, using covariance matrices. Non-parametric are various simulation methods, e.g. Monte Carlo simulations or calculations based on historical data. To read more about various methods how to calculate VaR see [14]. Nevertheless, the weakness results from the fact that different methods give different results for VaR. The methods are inconsistent even when the same method is used by different companies.

The fact that VaR does not provide information about the risk beyond the value it calculated, might be considered as its largest weakness. Intuitively, this drawback can be eliminated by replacing VaR with

---

[3]The situation, to which the hedge fund LTCM (Long Term Capital Management) got into in 1998 with losses of 4,6 billion USD, theoretically should not have happened not once in the existence of the Universe, or not even once in approx. 16 billion years.

the mean value of the losses beyond VaR. There exist two ways how to set the threshold for computing the expected value:

- Value of the loss. In this case, the expected value is calculated from all the losses that exceed the value of VaR, i.e. $-\mathbb{E}[X \mid X \leq -VaR(X)]$. This method was used to construct the risk measure TCE (Tail conditional expectation), also referred to as $TVaR$ (tail VaR).

- Probability of the loss. In this case, the expected value is calculated from all $(1-\alpha)\cdot 100\%$ worst-case scenarios, i.e. $-\mathbb{E}[X \mid X \in \{(1-\alpha)\cdot 100\% \text{ worst-case scenarios}\}]$. This is how risk measures ES (Expected shortfall) and $CVaR$ (Conditional VaR) were constructed.

The next section is devoted to $CVaR$.

3.3   Conditional value-at-risk

Conditional value-at-risk ($CVaR$) is another risk measure that, as stated above, does quantify the losses that might occur in the tail of the loss distribution. According to [22] $CVaR$ is defined as the expected value of losses that exceed the value of VaR, i.e. it is defined as the expected value of $(1 - \alpha) \cdot 100\%$ greatest losses. For normal distributions $CVaR$ provides results that are consistent with those of VaR and Markowitz's minimum variance, but $CVaR$ proves to be much more efficient when it comes to more difficult problems, e.g. dealing with discrete distributions.

Definition 3.9 ($CVaR^+$ and $CVaR^-$)**.** *Let $L \in \mathcal{L}$ be a loss function and $\alpha \in (0,1)$. The upper $CVaR_\alpha^+(\boldsymbol{x})$ and the lower[4] $CVaR_\alpha^-(\boldsymbol{x})$ of the loss associated with a decision $\boldsymbol{x} \in \mathcal{X}$ are defined as*

$$CVaR_\alpha^+(\boldsymbol{x}) = \mathbb{E}[L|L > VaR_\alpha(\boldsymbol{x})]$$

---

[4]$CVaR^-$ is sometimes referred to as tail-VaR [4].

$$CVaR_\alpha^-(\boldsymbol{x}) = \mathbb{E}[L|L \geq VaR_\alpha(\boldsymbol{x})].$$

Both these risk measures take the expected value of losses that are greater in value than VaR, i.e. they approach the problem of taking the expected value of $(1 - \alpha) \cdot 100\%$ greatest losses using the first method, see above. Constructing a risk measure from VaR this way keeps the properties of VaR and hence both $CVaR^+$ and $CVaR^-$ are not coherent. Generally the relationship $CVaR^- \leq CVaR^+$ holds, with equality holding only when the loss distribution function does not have a jump at point VaR_α.

When dealing with discrete distributions, it might be hard to properly define $(1 - \alpha) \cdot 100\%$ greatest losses. In this case, it is better to approach the problem of taking the expected value using α-tail, i.e. using the second method as stated above. This definition of $CVaR$ was first presented in [22] in reference to which we state the following definition.

Definition 3.10 ($CVaR$). *Let $L \in \mathcal{L}$ be a loss function with weights $\boldsymbol{x}$ and distribution function $\psi(\boldsymbol{x}, \xi)$, $\mathbb{E}[L] < +\infty$ and $\alpha \in (0, 1)$. Conditional value-at-risk associated with decision $\boldsymbol{x} \in \mathcal{X}$ at the confidence level α, $CVaR_\alpha$, is then defined as the mean of the α-tail distribution of L, which in turn is defined by*

$$\psi_\alpha(\boldsymbol{x}, \xi) = \begin{cases} 0 & \text{for } \xi < VaR_\alpha(\boldsymbol{x}), \\ \frac{\psi(\boldsymbol{x}, \xi) - \alpha}{1 - \alpha} & \text{for } \xi \geq VaR_\alpha(\boldsymbol{x}). \end{cases}$$

$CVaR$ can be calculated as a solution to a minimization problem described in the next theorem. This approach is particularly useful in solving optimization problems.

Theorem 3.1 ($CVaR$ minimization formula). *[23] Let $L = -\boldsymbol{x}^T \boldsymbol{\rho}$ be a loss function of a portfolio, $\alpha \in (0, 1)$. Then $CVaR_\alpha(\boldsymbol{x})$ is the*

solution to the following minimization problem:

$$CVaR_\alpha(\boldsymbol{x}) = \min_\xi \left\{ \xi + \frac{1}{1-\alpha} \mathbb{E}\{[L - \xi]^+\} \right\},$$

where $[v]^+ = \max\{0, v\}$. Possible solution to such a problem is any $\xi^* \in [VaR_\alpha(\boldsymbol{x}), VaR_\alpha^+(\boldsymbol{x})]$ or a single point $\xi^* = VaR_\alpha(\boldsymbol{x})$.

$CVaR$ attempts to eliminate some of the disadvantages of VaR, but it has drawbacks of its own. We will now state some of its pros and cons according to [25]. Let us again list the advantages first.

1. Interpretation. $CVaR$ has a clear economic interpretation. For example, if L is a random loss then the constraint $CVaR_\alpha(L) \leq L_0$ ensures that the average of $(1 - \alpha) \cdot 100\%$ greatest losses does not exceed the loss of previously chosen L_0.

2. Coherence. It is a coherent risk measure, see Definition 3.1 from [4].

3. Convexity. $CVaR$ of a convex combination of random variables $CVaR_\alpha(x_1 X_1 + \cdots + x_n X_n)$ is a convex function with respect to vector $\mathbf{x} \in \mathbb{R}^n$. So in financial settings, $CVaR$ of a portfolio is a convex function of the decision $\mathbf{x} \in \mathcal{X}$. This property is especially useful in the optimization of $CVaR$, which can then be reduced to convex programming and hence easily computed.

4. Continuity. $CVaR$ is continuous with respect to the confidence level α.

5. Probability. Similar to VaR, $CVaR$ also provides information about the probabilistic behaviour of the entire portfolio. Defining $CVaR_\alpha(X)$ for all confidence levels $\alpha \in (0, 1)$ completely specifies the distribution of random variable X. In this sense, it is superior to standard deviation.

When it comes to disadvantages of $CVaR$, we will follow [25] and state that the biggest drawback of $CVaR$ is the fact that it is sensitive

to estimation error, much more than VaR is. Accuracy of $CVaR$ is heavily affected by accuracy of tail modelling. In the absence of a good tail model, one should not count on $CVaR$. In financial settings, equally weighted portfolios may outperform $CVaR$-optimal portfolios out of sample when historical data have mean reverting characteristics.

This along with the fact that $CVaR$ is more difficult to back-test [5] might be the reason that $CVaR$ has not become a standard in finance industry, despite of its many useful mathematical properties. Even after the crisis it seems that regulators prefer to modify VaR and increase stress testing rather than to switch to $CVaR$. Nevertheless, as stated in [12] $CVaR$ is the most successful coherent risk measure so far and its popularity is rising.

When it comes to comparing the value of $CVaR$ with that of VaR both associated with a decision $\mathbf{x} \in \mathcal{X}$, for every $\alpha \in (0, 1)$ the relationship $VaR_\alpha(\mathbf{x}) \leq CVaR_\alpha(\mathbf{x})$ holds, with equality holding only when there is no chance of a loss greater than $VaR_\alpha(\mathbf{x})$.

When comparing stability of estimation of VaR and $CVaR$, appropriate confidence levels for VaR and $CVaR$ must be chosen in order to avoid comparison of VaR and $CVaR$ for the same level of α because they refer to different parts of the distribution.

$CVaR$ can actually be defined as a weighted average of VaR and $CVaR^+$ using the following theorem.

Theorem 3.2 (CVaR as a weighted average). *[23] Let $\lambda_\alpha(\boldsymbol{x}) \in [0, 1]$ be the probability assigned to the loss amount $VaR_\alpha(\boldsymbol{x})$ by the α-tail distribution defined in Definition (3.10), that is*

$$\lambda_\alpha(\boldsymbol{x}) = \frac{\psi(\boldsymbol{x}, VaR_\alpha(\boldsymbol{x})) - \alpha}{1 - \alpha}.$$

If $\psi(\boldsymbol{x}, VaR_\alpha(\boldsymbol{x})) < 1$, i.e. there is a chance of a loss exceeding $VaR_\alpha(\boldsymbol{x})$ ($\lambda_\alpha(\boldsymbol{x}) < 1$), then

$$CVaR_\alpha(\boldsymbol{x}) = \lambda_\alpha(\boldsymbol{x})VaR_\alpha(\boldsymbol{x}) + [1 - \lambda_\alpha(\boldsymbol{x})]CVaR_\alpha^+(\boldsymbol{x}).$$

Whereas if $\psi(\boldsymbol{x}, VaR_\alpha(\boldsymbol{x})) = 1$, i.e. $VaR_\alpha(\boldsymbol{x})$ is the highest loss that can occur ($\lambda_\alpha(\boldsymbol{x}) = 1$), then

$$CVaR_\alpha(\boldsymbol{x}) = VaR_\alpha(\boldsymbol{x}).$$

It might be surprising that we can get $CVaR$, a coherent and well behaved risk measure, as a weighted average of two risk measures, which are not.

In practice, the random vector of returns $\boldsymbol{\rho}$ has often discrete distribution. There are then S possible realizations of $\boldsymbol{\rho}$, which we will denote as $\boldsymbol{\rho}^s$, $s = 1, \ldots, S$. The loss L is then also a discrete variable that takes only finite number of values $l_s = -\mathbf{x}^T\boldsymbol{\rho}^s, s = 1, \ldots, S$. $CVaR$ works well in this discrete case and so we will now define $CVaR$ for a finite number of scenarios.

Definition 3.11 (CVaR and discete scenarios). *[30] Let the random vector of returns $\boldsymbol{\rho}$ have a discrete distribution. The distribution of a loss function L for a given $\boldsymbol{x} \in \mathcal{X}$ is also discrete, $P\{L = l_s\} = p_s$, $\sum_{s=1}^{S} p_s = 1$. Let s_α, $\alpha \in (0,1)$ be such an index for which*

$$\sum_{s=1}^{s_\alpha-1} p_s < \alpha \leq \sum_{s=1}^{s_\alpha} p_s$$

holds. Then

$$CVaR_\alpha(\boldsymbol{x}) = \frac{1}{1-\alpha}\left[\left(\sum_{s=1}^{s_\alpha} p_s - \alpha\right) s_\alpha + \sum_{s=s_\alpha+1}^{S} p_s l_s\right].$$

Analogy of $CVaR$ minimization formula also holds for discrete distributions.

Corollary 3.1. *[30] Let the random vector of returns $\boldsymbol{\rho}$ have a discrete distribution, with finite number of points $\boldsymbol{\rho}^s$ with probabilities p_s, $s =*

$1, \ldots, S$. *Then* $CVaR_\alpha(\boldsymbol{x})$ *can be calculated as*

$$CVaR_\alpha(\boldsymbol{x}) = \min_\xi \left\{ \xi + \frac{1}{1-\alpha} \sum_{s=1}^{S} p_s \left([-\boldsymbol{x}^T \boldsymbol{\rho}^s - \xi]^+ \right) \right\}.$$

In this chapter we introduced risk measures in general, defined variance, VaR and $CVaR$ and stated their advantages and disadvantages. For use in practical applications, we will formulate portfolio selection models for each of these risk measures in the following chapter.

Chapter 4

Portfolio selection models

As almost every investment is uncertain about the gain to be obtained in the future, a rational investor tries to reduce the risk related to investment. A naive solution to this risk minimization is diversification. A very old rule says that one should divide their wealth into three equal parts; one third to be put into deposits, one third to be invested into shares, and to buy gold for the remaining third [7]. More complex methods how to choose an optimal portfolio will be discussed in this chapter.

Although it is not necessary to make many mathematical assumptions when constructing an optimal portfolio, usually some reasonable economic conditions are assumed. Following [15] and adding comments about how realistic the assumptions are according to [7], we state the following assumptions:

1. An investor chooses portfolios with the highest expected return among those with the same risk (realistic for a rational investor).

2. An investor chooses portfolios with the smallest risk among those with the same expected return (realistic for a risk-averse investor).

3. The assets are infinitely divisible (limited, because trading on a stock exchange is usually performed in lots, for example, one hundred stocks, and it is either impossible to trade fractions of lots or there are extra costs imposed).

4. The investment horizon is one period in time (realistic, but multiple periods model might be better suited for practical applications).

5. There are no transaction costs and taxes (limited, but costs and taxes may partly be incorporated into the returns if they are linear functions of a traded volume).

6. There exists just one risk-free interest rate and all the investors can lend or borrow any amount of necessary funds at this rate (unrealistic, as interest rates for borrowing are usually higher than interest rates for lending).

7. All the assets in question are marketable (realistic).

8. *Short-sales* are not allowed (realistic; this is usually incorporated into legal regulations).

9. No investor can affect the returns of the respective assets substantially (restrictive, as it means that there is no investor with funds significantly exceeding the other investors' funds).

10. All necessary information is equally available to all the investors at the same time (restrictive, as efficient market hypothesis does not hold in reality and information asymmetry is present in financial markets).

When searching for an optimal portfolio using mean-risk models we are actually looking for an efficient portfolio, which we will now define according to [7].

Definition 4.1. *Let R (R^*) be an expected return of a portfolio as defined in Definition (2.1) and $\mathcal{R}(L)$ ($\mathcal{R}(L^*)$) be a risk associated with decision $\boldsymbol{x} \in \mathcal{X}$ ($\boldsymbol{x}^* \in \mathcal{X}$). Then portfolio $\boldsymbol{x}^*$ is called efficient portfolio if there is no other portfolio $\boldsymbol{x}$ such that*

$$(R^* < R \wedge \mathcal{R}(L^*) \geq \mathcal{R}(L)) \vee (R^* = R \wedge \mathcal{R}(L^*) > \mathcal{R}(L)).$$

In other words, an efficient portfolio is a portfolio for which there is no other portfolio with greater expected return and smaller or equal risk or with equal expected return and smaller risk.

According to [7] mean-risk models can be formulated in three equivalent ways without further assumptions as follows.

$$\max_{\mathbf{x} \in \mathcal{X}} \quad \mathbb{E}(Y) - \lambda \mathcal{R}(L) \tag{4.1}$$

$$\begin{aligned} \max_{\mathbf{x} \in \mathcal{X}} \quad & \mathbb{E}(Y) \\ \text{s.t.} \quad & \mathcal{R}(L) \leq q_{\max} \end{aligned} \tag{4.2}$$

$$\begin{aligned} \min_{\mathbf{x} \in \mathcal{X}} \quad & \mathcal{R}(L) \\ \text{s.t.} \quad & \mathbb{E}(Y) \geq r_{\min} \end{aligned} \tag{4.3}$$

The first formulation of a model (4.1) represents the problem of maximization of the expected return of a portfolio adjusted for an associated risk. The value of parameter $\lambda \geq 0$ is related to the investor's risk aversion (large values of λ correspond to a risk averse investor, while small values of λ are typical for risk loving investors). The second formulation (4.2) shows maximization of the expected return of a portfolio subject to the highest value of risk $q_{\max}$ an investor is willing to undertake. The third formulation (4.3), on the other hand, presents the minimization of the risk subject to a minimal acceptable return. This last formulation is the most natural, as an investor usually has an idea about the minimal return they require, while risk might be assessed more subjectively and hence, we will use this model when formulating mean-risk models.

It is stated in [27] that if $\mathbf{x}^*(\lambda)$ solves the problem formulated as in (4.1) for a fixed value of λ, it also solves the minimization problem shown in (4.3) for $r_{\min} = R^*$, where $R^* = \mathbb{E}\boldsymbol{\rho}^T \mathbf{x}^*(\lambda)$ is the expected return of a portfolio associated with decision $\mathbf{x}^*(\lambda)$.

4.1 Mean-Variance Model

According to [1], portfolio $\mathbf{x}^*$ is mean-variance efficient when $\mathbf{x}^*$ solves the following minimization problem:

$$\min_{\mathbf{x} \in \mathcal{X}} \quad \sigma^2(\mathbf{x})$$
$$\text{s.t.} \quad \mathbb{E}\{\mathbf{x}^T \boldsymbol{\rho}\} \geq r_{\min} \tag{4.4}$$
$$\sum_{i=0}^{N} x_i = 1$$

where $r_{\min}$ is the minimal required expected return of a portfolio. This is a general formulation of Markowitz model. We will now state a specific formulation of this model for finite number of scenarios.

$$\min_{\mathbf{x} \in \mathcal{X}} \quad \sigma^2(\mathbf{x})$$
$$\text{s.t.} \quad \frac{1}{S}\sum_{s=1}^{S} \mathbf{x}^T \boldsymbol{\rho}^s \geq r_{\min} \tag{4.5}$$
$$\sum_{i=0}^{N} x_i = 1,$$

where $\mathcal{X}$ is as stated above, i.e. $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^{N+1} \mid x_i \geq 0\}$, $i = 0, 1, \ldots, N$.

Because variance as a measure of risk is easy to obtain and interpret, this model is quite intuitive and easy to use. Variance, however, has its drawbacks, stated in the section devoted to variance and standard deviation, that limit this model. In order to deal with these shortcomings and to perform better in practical applications, mean-risk models using different risk measures were introduced.

4.2 Mean-VaR Model

We will now formulate the portfolio optimization problem using VaR as the risk measure in reference to [19]. The mean-VaR opti-

mization problem is of the following form:

$$\min_{\mathbf{x} \in \mathcal{X}} \quad VaR_\alpha(\mathbf{x})$$
$$\text{s.t.} \quad \mathbb{E}\{\mathbf{x}^T \boldsymbol{\rho}\} \geq r_{\min}$$
$$\sum_{i=0}^{N} x_i = 1. \tag{4.6}$$

For practical applications, when L is a discrete random variable, we formulate the mean-VaR model for finite number of scenarios.

Let us define the function $M_{[k:N]}(u^1, \ldots, u^N)$ to denote the k-th largest value among $u^1, \ldots, u^N$. We can see that $M_{[1:N]}$ represents the minimum and $M_{[N:N]}$ the maximum. According to [19], VaR for the discrete distribution is calculated to be

$$VaR_\alpha(\mathbf{x}) = M_{[\lfloor \alpha S \rfloor : S]}(-\mathbf{x}^T \boldsymbol{\rho}^1, \ldots, -\mathbf{x}^T \boldsymbol{\rho}^S),$$

where $\lfloor x \rfloor$ represents the floor of the number $x \in \mathbb{R}$.

The discrete portfolio optimization problem for VaR is then formulated as follows.

$$\min_{\mathbf{x} \in \mathcal{X}} \quad M_{[\lfloor \alpha S \rfloor : S]}(-\mathbf{x}^T \boldsymbol{\rho}^1, \ldots, -\mathbf{x}^T \boldsymbol{\rho}^S)$$
$$\text{s.t.} \quad \frac{1}{S} \sum_{s=1}^{S} \mathbf{x}^T \boldsymbol{\rho}^s \geq r_{\min}$$
$$\sum_{i=0}^{N} x_i = 1, \tag{4.7}$$

which can, in reference to [30], be reformulated to:

$$
\begin{aligned}
\min_{\nu, \mathbf{x}, \delta^s} \quad & \nu \\
\text{s.t.} \quad & -\mathbf{x}^T \boldsymbol{\rho}^s \leq \nu + M\delta^s, s = 1, 2, \ldots, S \\
& \sum_{s=1}^{S} \delta^s = \lfloor (1-\alpha)S \rfloor \\
& \delta^s \in \{0; 1\}, \ s = 1, 2, \ldots, S \\
& \frac{1}{S} \sum_{s=1}^{S} \mathbf{x}^T \boldsymbol{\rho}^s \geq r_{\min} \\
& \sum_{i=0}^{N} x_i = 1
\end{aligned}
\tag{4.8}
$$

where $\mathbf{x} \in \mathcal{X}$ and $\mathcal{X}$ is as above and M is a constant large enough, so that $M \geq \max_{i,s} \rho_i^s - \min_{i,s} \rho_i^s$.

4.3 Mean-CVaR Model

We will formulate the mean-$CVaR$ optimization problem using Theorem 3.1 for the definition of $CVaR$.

$$
\begin{aligned}
\min_{\mathbf{x}, \xi} \quad & \xi + \frac{1}{1-\alpha} \mathbb{E}\{Z\} \\
\text{s.t.} \quad & Z \geq -\mathbf{x}^T \boldsymbol{\rho} - \xi \\
& Z \geq 0 \\
& \mathbb{E}\{\mathbf{x}^T \boldsymbol{\rho}\} \geq r_{\min} \\
& \sum_{i=0}^{N} x_i = 1.
\end{aligned}
\tag{4.9}
$$

For this problem, every local minimum is also a global minimum, which is a big advantage of the $CVaR$ risk measure over the VaR risk measure.

According to [30], for $\boldsymbol{\rho}$ with discrete distribution, when all scenar-

ios are equally probable, with probability $p_s = \frac{1}{S}$, $s = 1, \ldots, S$, the optimization problem can be reformulated as follows.

$$
\begin{aligned}
\min_{\mathbf{x}, \xi, z} \quad & \xi + \frac{1}{(1-\alpha)S} \sum_{s=1}^{S} z^s \\
\text{s.t.} \quad & z^s \geq -\mathbf{x}^T \boldsymbol{\rho}^s - \xi \\
& z^s \geq 0, \ s = 1, \ldots, S \\
& \frac{1}{S} \sum_{s=1}^{S} \mathbf{x}^T \boldsymbol{\rho}^s \geq r_{\min} \\
& \sum_{i=0}^{N} x_i = 1,
\end{aligned}
\tag{4.10}
$$

where $\mathbf{x} \in \mathcal{X}$ and $\mathcal{X}$ as above.

For now we have formulated classical mean-risk models with variance, VaR and $CVaR$ as risk measures in general and then specifically for discrete distributions. All of these models are, however, very sensitive to input parameters [8] and hence new robust approaches were taken. We will focus on this robust concept in the next chapter.

Chapter 5

Robust portfolio selection models

The above formulated portfolio selection models are very sensitive to input parameters, which, however, are never known for certain and hence have to be estimated. This causes another risk that arises from the estimation.

Robust optimization can incorporate this estimation risk into the decision making process in portfolio choice. Generally speaking, robust optimization refers to finding solutions to given optimization problems with uncertain input parameters that will achieve good objective values for all, or most, realizations of the uncertain input parameters. It should be noted, however, that there are different interpretations of robustness that lead to different mathematical formulations [8]. Here, we take the pessimistic view of robustness and look for a solution that has the best performance in the worst possible scenario (worst-case), i.e. the maximum risk is being minimized and/or the minimal possible return is being maximized.

Given a problem with uncertain inputs and an uncertainty set for these inputs, our robust optimization approach addresses the following problem: What choice of the variables of the problem will optimize the worst case objective value? That is, for each choice of the decision variables, we consider the worst case realization of the data and evaluate the corresponding objective value, and then pick the set of values for the variables with the best worst-case objective.

5.1 Worst-case Mean-Variance Model

Despite the profound theoretical concept of the mean-variance model developed by Markowitz in 1952 [15], this model is treated sceptically in the investment practice. One of the reasons is the counter-intuitive behaviour of the optimal portfolios computed by this approach. Optimal portfolios tend to concentrate on a small subset of the available securities, and seem not to be well diversified. Furthermore, optimal portfolios are often sensitive to changes in the input parameters of the problem (expected returns and the covariance matrix). These input parameters cannot be observed and we have to rely on the estimates.

Hence, the inputs to the mean-variance model need to be very accurately estimated. This is, however, very difficult, especially in the case of expected return estimation. The use of estimates in the mean-variance model introduces an estimation risk in portfolio choice, and methods for optimal selection of portfolios must take this risk into account.

The uncertainty in the estimates can be described with the help of uncertainty sets, which include all, or most, possible realizations of the uncertain input parameters. The actual specification of proper uncertainty set is essentially the hardest part in robust portfolio selection. The worst-case approach to robust portfolio selection described in this section can be significantly influenced by outliers in the data and so uncertainty sets need to be carefully chosen. Because of the outlier values uncertainty sets usually include rather most, but not necessarily all possible realizations of the uncertain input parameters. There are different methods how to generate these uncertainty sets. Some are based on estimates from a factor model [11], some are simply a collection of estimates for the unknown input parameters or several possible scenarios, as in [24]. Perhaps, the simplest and hence, most often used are box and ellipsoidal uncertainty sets. In what follows, we will deal only with box uncertainty sets, i.e. setting an interval

for our estimate instead of using only one, often incorrect value. The end-points of the interval may be set, for example, as the minimum and maximum values of the corresponding statistic calculated from historical data or from simulated scenarios, or taken as reasonable estimates.

In the case of mean-variance optimization problems, robustness can be achieved, following [27], using uncertainty set for the expected return $\mathbf{r}$ and the covariance matrix $\mathbf{V}$. We will use box uncertainty sets, in which case we can write:

$$
\begin{aligned}
U_r &= \{\mathbf{r} : \mathbf{r}^L \leq \mathbf{r} \leq \mathbf{r}^U\} \\
U_V &= \{\mathbf{V} : \mathbf{V}^L \leq \mathbf{V} \leq \mathbf{V}^U; \mathbf{V} \succeq 0\} \\
U &= \{(\mathbf{r}, \mathbf{V}) : \mathbf{r} \in U_r, \mathbf{V} \in U_V\}.
\end{aligned}
$$

The restriction $\mathbf{V} \succeq 0$ indicates that $\mathbf{V}$ is a symmetric positive semidefinite matrix, which is a necessary property of the uncertain $\mathbf{V}$ to be a covariance matrix. Inequality $\mathbf{V} \leq \mathbf{V}^U$ means that every element in matrix $\mathbf{V}^U$ is greater than or equal to the corresponding element in matrix $\mathbf{V}$. It holds analogically for the inequality $\mathbf{V}^L \leq \mathbf{V}$.

Let us restate that we consider decision vectors that belong to the following set $\mathcal{X} = \left\{\mathbf{x} \in \mathbb{R}^{N+1} \mid x_i \geq 0\right\}, i = 0, 1, \ldots, N$. Given the uncertainty set U, the robust version of the mean-variance optimization problem can be expressed as follows.

$$
\begin{aligned}
\min_{\mathbf{x} \in \mathcal{X}} \ \max_{\mathbf{V} \in U_V} \ & \mathbf{x}^T \mathbf{V} \mathbf{x} \\
\text{s.t.} \quad \min_{\mathbf{r} \in U_r} \ & \mathbf{r}^T \mathbf{x} \geq r_{\min} \\
& \sum_{i=0}^{N} x_i = 1.
\end{aligned}
\tag{5.1}
$$

Let us further assume that the uncertainty set U is such that the matrix $\mathbf{V}^U$ is positive semidefinite, i.e. an upper bound on the covariance matrix forms an acceptable covariance matrix itself. For a

general case, when these simplifying assumptions are relaxed, we refer interested readers to [27]. In this case, it is proven in [27] that the worst-case realization of the data is the same regardless of what portfolio is chosen. Specifically, expected returns are realized at their lowest possible values and the covariances are realized at their highest possible values.

Hence, the robust mean-variance portfolio selection reduces to the following minimization problem.

$$
\begin{aligned}
\min_{\mathbf{x} \in \mathcal{X}} \quad & \mathbf{x}^T \mathbf{V}^U \mathbf{x} \\
\text{s. t.} \quad & \left(\mathbf{r}^L\right)^T \mathbf{x} \geq r_{\min} \\
& \sum_{i=0}^{N} x_i = 1.
\end{aligned}
\tag{5.2}
$$

As is shown in [6], errors in means are about ten times as important as errors in variances and covariances and errors in variances are about twice as important as errors in covariances. Hence, we will assume that covariances are given and we will take into account uncertainties only in the estimates of expected return and variance. We can consider uncertainties in the estimates of both the variance and the expected return, but we can also allow uncertainty only in one of them, either in the estimate of variance or the expected return. Let us now consider these three possibilities and formulate corresponding optimization problems in more detail.

First, when only uncertainty in the estimate of expected return is

considered, we get the following problem.

$$\min_{\mathbf{x} \in \mathcal{X}} \quad \mathbf{x}^T \mathbf{V} \mathbf{x}$$

$$\text{s.t.} \quad \frac{1}{S} \sum_{s=1}^{S} \mathbf{x}^T \mathbf{r}_s^L \geq r_{\min}, \tag{5.3}$$

$$\sum_{i=0}^{N} x_i = 1.$$

We can rewrite the constraint for mean calculated from different scenarios as follows: $\frac{1}{S} \sum_{s=1}^{S} \sum_{i=0}^{N} (r_i - \epsilon_i) x_i \geq r_{\min}$, where $\epsilon_i > 0$ are chosen so that for the elements of $\mathbf{r}^L$, it holds that $r_i^L = r_i - \epsilon_i$. For each of the N assets we consider different uncertainty, hence ϵ_i differs for different $i = 0, 1, \ldots, N$. Had we considered that $\epsilon_i = \epsilon_j$ for all $i \neq j$, $i, j = 0, 1, \ldots, N$, we would get the original mean-variance problem, where the required minimal return $r_{\min}$ would be reduced by ϵ.

Next, considering uncertainty only in the estimate for variance, we can write the problem as follows.

$$\min_{\mathbf{x} \in \mathcal{X}} \quad \sum_{i=0}^{N} \sum_{j \neq i}^{N} x_i x_j cov(\rho_i, \rho_j) + \sum_{i=0}^{N} x_i^2 (var(\rho_i) + \delta_i)$$

$$\text{s.t.} \quad \frac{1}{S} \sum_{s=1}^{S} \mathbf{x}^T \mathbf{r}_s^L \geq r_{\min} \tag{5.4}$$

$$\sum_{i=0}^{N} x_i = 1,$$

where $\delta_i > 0$ is similarly set so that for diagonal elements of $\mathbf{V}^U$, i.e. variances, it holds that $var(\rho_i)^U = var(\rho_i) + \delta_i$.

Finally, considering uncertainty both in the expected return and in

the variance, we can rewrite formulation of 5.2 as follows.

$$\min_{\mathbf{x} \in \mathcal{X}} \quad \sum_{i=0}^{N} \sum_{j \neq i}^{N} x_i x_j cov(\rho_i, \rho_j) + \sum_{i=0}^{N} x_i^2 (var(\rho_i) + \delta_i)$$

$$\text{s.t.} \quad \frac{1}{S} \sum_{s=1}^{S} \sum_{i=0}^{N} (r_i - \epsilon_i) x_i \geq r_{\min} \quad (5.5)$$

$$\sum_{i=0}^{N} x_i = 1,$$

where $\delta_i > 0$ and $\epsilon_i > 0$ are set as above.

5.2   Worst-case Mean-VaR Model

In general, however, we can formulate the robust portfolio selection problem as follows:

$$\min_{\mathbf{x} \in \mathcal{X}} \sup_{p \in \mathcal{P}} \quad \mathcal{R}_p(L)$$

$$\text{s.t.} \quad \mathbb{E}_p(Y) \geq r_{\min} \quad \forall p \in \mathcal{P}, \quad (5.6)$$

where $\mathcal{P}$ stands for the set of probability distributions considered, $\mathcal{R}_p(L)$ and $\mathbb{E}_p(Y)$ represent the risk and expected return under probability distribution p.

Definition 5.1. *Let $L = -\boldsymbol{x}^T \boldsymbol{\rho}$ be the loss function, where $\boldsymbol{x}$ is the investor's decision and $\boldsymbol{\rho}$ is a random vector with probability distribution $p \in \mathcal{P}$. Let $\mathcal{R} : \mathcal{L} \to \mathbb{R}$ is a measure of risk. Then we define the worst-case risk measure $\mathcal{R}^{WC}$ to the risk measure $\mathcal{R}$ as:*

$$\mathcal{R}^{RW}(L) = \sup_{p \in \mathcal{P}} \mathcal{R}(L).$$

Possible choices for $\mathcal{P}$ are listed in [28] and references therein. For example, $\mathcal{P}$ can be chosen so that it contains only probability distributions which fulfil certain moment conditions, belong to a neighbourhood of some hypothetical probability distribution or possess some

additional information. For discrete distributions one can also model the uncertainty in various ways: as incomplete information in the values of scenarios, incomplete information in probabilities or a combination of these two. Formulation of this last approach can be found in [30].

Here we will state only the case, when $\mathcal{P}$ contains all possible probability distributions of $\boldsymbol{\rho}$ that fulfil some moment conditions. Especially, we will assume that the set $\mathcal{P}$ contains all probability distributions that are given by the same first two moments. Hence, for given $\mathbf{r}_\rho$ and $\mathbf{V}_\rho$, we get:

$$\mathcal{P}^M = \{p : \mathbb{E}(\boldsymbol{\rho}) = \mathbf{r}_\rho, \mathbb{E}(\boldsymbol{\rho} - \mathbf{r}_\rho)^2 = \mathbf{V}_\rho\}.$$

For this specific case we will first formulate the worst-case VaR model and later show, according to [28], that in this special case worst-case VaR and worst-case $CVaR$ coincide.

Let us first state the proposition for one-sided Chebyshev bound, which will be of a use in formulating the worst-case VaR.

Proposition 5.1 (one-sided Chebyshev bound). *[30] Let $L = -\boldsymbol{x}^T\boldsymbol{\rho}$ be a loss function with the expected value equal to $\boldsymbol{r}_\rho^T\boldsymbol{x}$ and the variance of $\boldsymbol{x}^T\boldsymbol{V}_\rho\boldsymbol{x}$. Then the upper bound for the distribution function $\psi(\boldsymbol{x}, \xi) = P(L \leq \xi)$, is given by:*

$$\sup_{p \in \mathcal{P}^M} P(L \leq \xi) = \begin{cases} \frac{\boldsymbol{x}^T\boldsymbol{V}_\rho\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{V}_\rho\boldsymbol{x} + (\boldsymbol{r}_\rho^T\boldsymbol{x} - \xi)^2} & \textit{for } \xi < \boldsymbol{r}_\rho^T\boldsymbol{x}, \\ 1 & \textit{for } \xi \geq \boldsymbol{r}_\rho^T\boldsymbol{x}. \end{cases}$$

With the above stated Proposition 5.1 and formulation (5.6) of general robust portfolio selection problem, we can formulate the robust

mean-VaR problem for $\xi < \mathbf{r}_\rho^T \mathbf{x}$ as follows:

$$\min_{\mathbf{x} \in \mathcal{X}} \quad \mathbf{r}_\rho^T \mathbf{x} + \kappa(\alpha)\sqrt{\mathbf{x}^T \mathbf{V}_\rho \mathbf{x}}$$
$$\text{s.t.} \quad \mathbb{E}(\mathbf{x}^T \boldsymbol{\rho}) \geq r_{\min} \tag{5.7}$$
$$\sum_{i=0}^{N} x_i = 1,$$

where $\kappa(\alpha) = \sqrt{\frac{\alpha}{1-\alpha}}$ and $\mathcal{X}$ is as above, i.e. $\mathcal{X} = \left\{ \mathbf{x} \in \mathbb{R}^{N+1} \mid x_i \geq 0 \right\}$, $i = 0, 1, \dots, N$.

5.3 Worst-case Mean-CVaR Model

According to [28], we can now state the worst-case conditional value-at-risk as follows:

$$CVaR_{\alpha,\mathcal{P}}^{WC}(\mathbf{x}) = \sup_{p \in \mathcal{P}} \min_{\xi \in \mathbb{R}} \left\{ \xi + \frac{1}{1-\alpha}\mathbb{E}_p([L - \xi]^+) \right\}.$$

For a given expected value $\mathbf{r}_\rho$ and variance $\mathbf{V}_\rho$ we can rewrite this definition of $CVaR_{\alpha,\mathcal{P}}^{WC}(\mathbf{x})$ in successive steps, in reference to [28]. First, due to the finite expected value and convexity of the inner objective function we can interchange maximum and minimum. Then we can write

$$\max_{p \in \mathcal{P}^M} \mathbb{E}_p\{[L - \xi]^+\} = \frac{1}{2}\left[\mathbf{r}_\rho^T \mathbf{x} - \xi + \sqrt{\mathbf{x}^T \mathbf{V} \mathbf{x} + (\mathbf{r}_\rho^T \mathbf{x} - \xi)^2} \right].$$

Then by solving the problem

$$\min_{\xi \in \mathbb{R}} \xi + \frac{1}{2(1-\alpha)}\left[\mathbf{r}_\rho^T \mathbf{x} - \xi + \sqrt{\mathbf{x}^T \mathbf{V} \mathbf{x} + (\mathbf{r}_\rho^T \mathbf{x} - \xi)^2} \right]$$

we obtain for $\alpha \in [\frac{1}{2}, 1)$ and $\xi < \mathbf{r}_\rho^T \mathbf{x}$, the optimal solution

$$\xi^* = \mathbf{r}_\rho^T \mathbf{x} + \frac{1 - 2(1-\alpha)}{2\sqrt{\alpha(1-\alpha)}}\sqrt{\mathbf{x}^T \mathbf{V} \mathbf{x}}$$

and then the formula for $CVaR^{WC}_{\alpha,\mathcal{P}}(\mathbf{x})$ can be rewritten as follows:

$$CVaR^{WC}_{\alpha,\mathcal{P}}(\mathbf{x}) = \mathbf{r}_\rho^T \mathbf{x} + \sqrt{\frac{\alpha}{1-\alpha}} \sqrt{\mathbf{x}^T \mathbf{V} \mathbf{x}}.$$

Setting $\kappa(\alpha) = \sqrt{\frac{\alpha}{1-\alpha}}$, we see that for a probability distribution identified by its first two moments and for $\xi < \mathbf{r}_\rho^T \mathbf{x}$, the robust model for $CVaR_{\alpha,\mathcal{P}}{}^{WC}(\mathbf{x})$ coincides with that for $VaR_{\alpha,\mathcal{P}}{}^{WC}(\mathbf{x})$ formulated in (5.7).

Chapter 6

Numerical example

In this section we would like to show the idea of robustness by comparing simple mean-variance model (4.5) with its three robust counterparts (5.3), (5.4) and (5.5) applied to real historical data. We solved the optimization problems using free software for statistical computing R [21]. Specifically, its solver quadprog [26] was used for solving quadratic programming problems. The source code used for solving the optimization problems can be found in Appendix C. Please note that for generating the figures found in this chapter, the code was run in the following way:

```
eee.var <- diag(cov(returns))
eee <- colMeans(rets$R) * 0.2
eff4 <- eff.mean.variance(returns, e=eee, e.var=eee.var,
    robust.mean=TRUE, robust.var=TRUE)
```

Where `returns` are shown as a matrix of historical returns for each asset obtained from Yahoo Finance. The parameters `e` and `e.var` denote ϵ_i and δ_i in models (5.3), (5.4) and (5.5). The procedure returned optimal asset allocations for feasible minimal required returns, which were then plotted separately.

Source for the data was financial portal Yahoo Finance [29]. We used historical prices of shares of 12 selected companies listed on New York Stock Exchange, NYSE, which is the world's largest stock exchange by market capitalization. Companies were selected from various industries, so that diversification can work well. The list of the

companies along with their ticker symbol can be found in Table 6.1. We give a short description of each company in Appendix A. All shares are USD denominated and graphs of their price developments are presented in Appendix B.

Company name	Ticker	Company name	Ticker
Barnes & Noble, Inc.	BKS	Hewlett-Packard Company	HPQ
The Boeing Company	BA	McDonald's Corporation	MCD
Credit Suisse Group AG	CS	PepsiCo, Inc.	PEP
Exxon Mobil Corporation	XOM	Pfizer Inc.	PFE
Ford Motor Company	F	Wal-Mart Stores, Inc.	WMT
General Electric Company	GE	The Walt Disney Company	DIS

Table 6.1: Selected companies

In order to calculate historical returns of shares we used close prices adjusted for dividends and splits. These close prices were taken at the end of the first trading day of each week in the period from 26.4.2004 until 27.4.2014, which gave us 521 scenarios for weekly returns. These weekly returns were calculated from the following formula:

$$r_{s,i} = \frac{P_{s,i}^{Adj} - P_{s-1,i}^{Adj}}{P_{s-1,i}^{Adj}}, \ s = 1, \ldots, S, \ i = 0, 1, \ldots, N,$$

where $P_{s,i}^{Adj}$ is an adjusted close price of a share of the i-th company at the end of the first trading day in the s-th week. The average weekly returns (in %) for our 12 companies can be summarized as follows:

BKS	BA	CS	XOM	F	GE	HPQ	MCD	PEP	PFE	WMT	DIS
0.274	0.339	0.197	0.252	0.310	0.130	0.221	0.337	0.159	0.102	0.132	0.326

We can see that the highest positive returns were achieved by shares of The Boeing Company and Mc Donald's Corporation. The least paying were shares of Pfizer Inc. and General Electric Company.

For the risk-free asset we used 6 month US treasury zero coupon bond. We got the data for its annual historical yields calculated weekly from [9]. Since the yields of US treasury bonds fell significantly after the financial crisis 2007-2008 and we assume these low yields will

prevail in the considerable future, we simulated the 521 observations for risk-free asset from normal distribution with mean of 0.16% and standard deviation of 0.09%. The values mean and standard deviation were calculated from historical annual yields calculated weekly (every Friday) from the period 28.11.2008 until 2.5.2014. This simulation of 521 risk-free scenarios had to be done in order to fill the covariance matrix $\mathbf{V}$. Although theoretically it is enough for $\mathbf{V}$ to be positive semidefinite, see above, the solver used required $\mathbf{V}$ to be positive definite. This prevented us from using theoretical risk-free asset, which would really have its risk and return set to 0.

Since we consider neither *short-selling* nor borrowing of money, we can assume that an investor chooses the minimal required weekly return in the interval [0.003%; 0.339%] with corresponding annual returns in the interval [0.16%; 17.628%]. The lower bound is associated with weekly (annual) risk-free return and the upper bound with weekly (annual) return of the highest paying asset, i.e. shares of The Boeing Company.
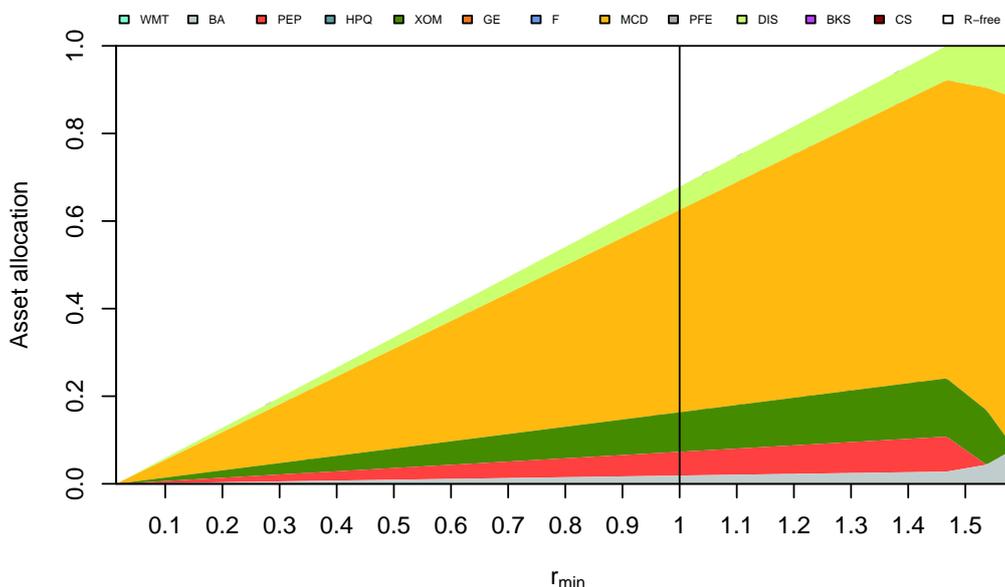
Figure 6.1: Asset allocation using mean-variance model

Solving the problem presented in the classical Markowitz mean-variance model (4.5) for the historical data, gives us the asset alloca-

tion shown in Figure 6.1.

We can see that a rational investor invests mainly into the shares of Mc Donald's Corporation, Exxon Mobile Corporation, PepsiCo Inc. and The Walt Disney Company; small portion of investor's wealth is allocated into other shares and the rest is invested into the risk-free asset.

Until the minimal required return (r_{min}) reaches the value of 1.47, the proportion of different companies' shares in the risky part of the portfolio does not change with increasing minimal required return (r_{min}). The only thing changing is the proportion of the risky part of the portfolio to the risk-free asset in the entire portfolio. When the minimal required return reaches 1.47 and goes above, all the wealth is allocated into risky assets and their proportion in the portfolio is also changing, in order to fulfil the requirement of higher minimal return.
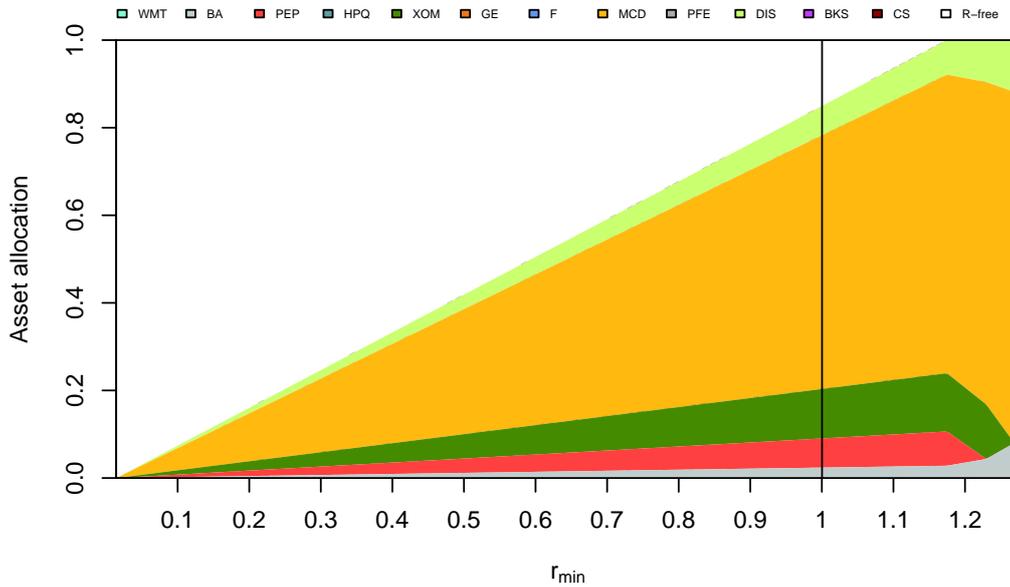
Figure 6.2: Asset allocation with interval for mean

Now we show what will change, when allowing uncertainty for the mean. So we will now compute the problem described in model 5.3 with our historical data. We assume that the worst returns $\mathbf{r}^L$ are such that $r_i^L = 0.8r_i$, i.e. original returns fell by 20% and hence we set $\epsilon_i = 0.2r_i$. The solution to such a problem is represented in Figure 6.2.

We can see that an investor still invests mainly into the shares of Mc Donald's Corporation, Exxon Mobile Corporation, PepsiCo Inc. and The Walt Disney Company; partly into other shares and invests the rest of their wealth into the risk-free asset. Comparing Figure 6.1 with Figure 6.2, we can see that for a given required minimal return, the proportion of risky higher-yielding assets in the entire portfolio has risen. This is due to the fact that assuming uncertainty in the estimated mean return decreases the return for any given level of risk.

Further, we can assume uncertainty only in the estimate of variance, we will show a solution to model 5.4. In this case we assume that a situation worsens so much that $\mathbf{V}^U$ is such that $var(\rho_i)^U = 2var(\rho_i)$, i.e. that the risk increases two times and hence $\delta_i = var(\rho_i)$. The solution to such a problem is presented in Figure 6.3.
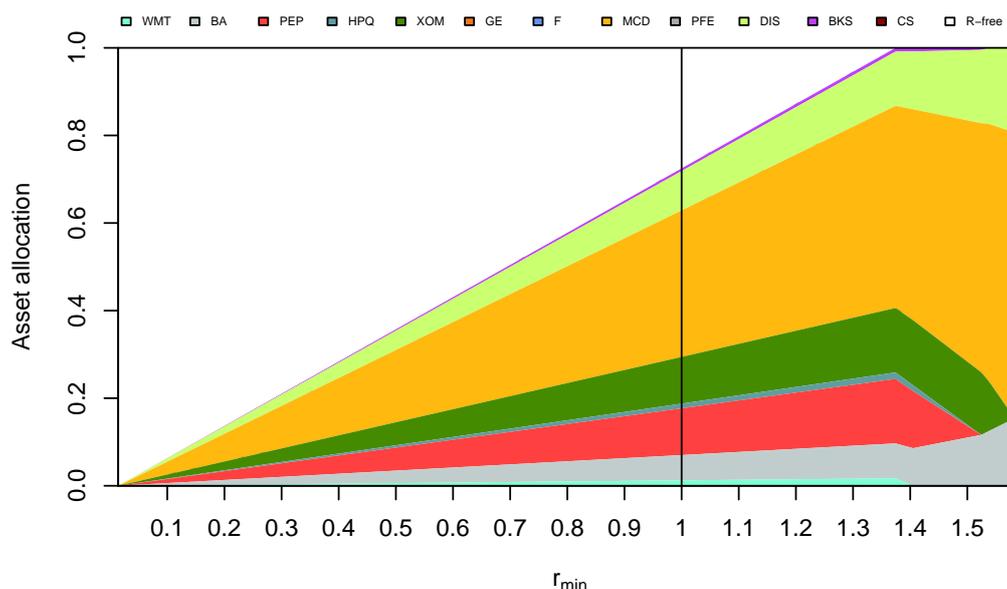
Figure 6.3: Asset allocation with interval for variance

We can see that considering uncertainty in the estimates for variance, resulted in much more diversified portfolio that the one shown in Figure 6.1 and so the portfolio is more resistant to the potential risk.

Finally, considering uncertainty both in the estimate for mean and variance, i.e. finding solution to the problem represented in model 5.5,

we get the optimal asset allocation as shown in Figure 6.4.
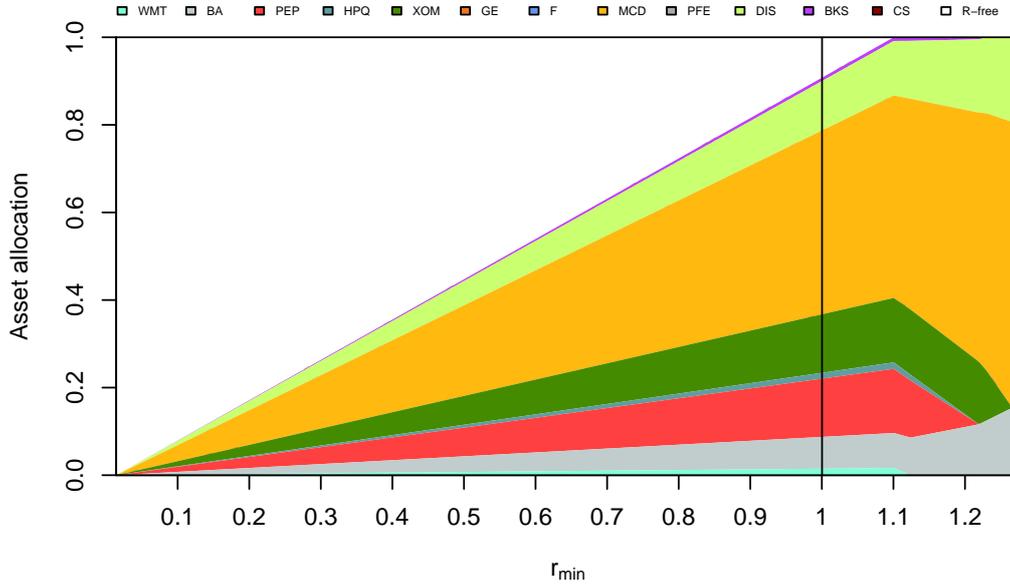
Figure 6.4: Asset allocation with intervals for both mean and variance

We can see the portfolio is now more diversified with lower portion of risk-free asset compared to the one shown in Figure 6.1. Moreover, the proportion of risky assets in the portfolio is changing with minimal required returns (r_{min}) above 1.1. Highest r_{min} now require investing only into three risky assets, i.e. shares of Mc Donald's Corporation, The Boeing Company and The Walt Disney Company.

For the above stated models and using the same values for ϵ_i and δ_i, i.e. $\epsilon_i = 0.2r_i$ and $\delta_i = var_i$, we can also show the results in the form of efficient frontiers as is shown in the following figure.
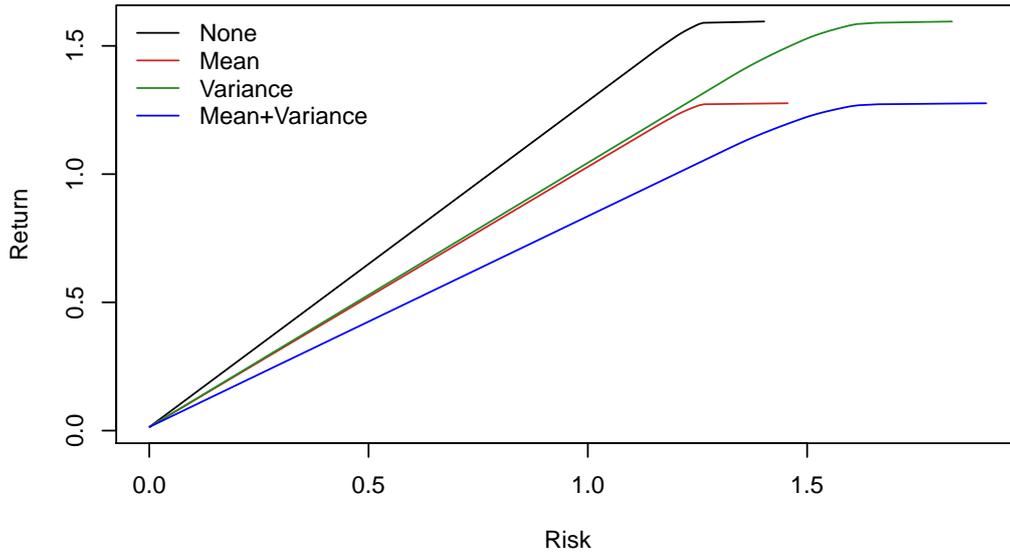
Figure 6.5: Efficient frontiers

For any level of risk, measured by variance, Figure 6.5 shows the associated return of a portfolio for each of the discussed models. Individual efficient frontiers end at the point when maximal possible return for each the models is reached and hence, higher risk taken does not result in higher return.

The robust counterparts weigh the riskiness of individual asset returns more and hence for any level of risk, these robust problems give lower value of expected return than their non-robust alternative.

As stated in Section 5.1, one of the hardest parts in robust portfolio selection is to properly identify the uncertainty set. In our example we set the uncertainty set rather pessimistically. Setting the values of ϵ_i and δ_i differently, gave us slightly different results. The lower ϵ_i and δ_i were set, the closer were the robust solutions to the classical non-robust mean-variance solution.

Chapter 7

Conclusion

In this thesis we took the mean-risk approach to portfolio optimization. We defined general risk measures and introduced three commonly used ones: variance, VaR and $CVaR$. For each of these risk measures we formulated corresponding mean-risk models. All of these models rely on some estimate for input parameters, and hence estimation error is present. The uncertainty in data and in estimates used for optimization models should not be ignored. Results should be robust enough, so that small changes in input parameters do not result in extreme changes in the solutions. Robust approaches do not ignore this estimation error and incorporate it directly into the calculation of optimal portfolios. We focused on specifying robust versions to classical mean-variance model, which we also applied to historical data and compared the results with non-robust Markowitz's model.

All of the problems were computed as quadratic optimization problems using quadprog solver in free statistical software R. We found that in comparison to non-robust Markowitz's model robust versions resulted in lower expected returns when an interval for the estimate of the mean was imposed and in more diversified portfolios when we assumed uncertainty in the estimate of variance. Generally our results suggest that robust versions work to make the resulting optimal portfolios less sensitive to input parameters, which is in accordance with the results in [27]. For the other two risk measures we only stated the

special case, when the first two moments are known.

For further study we leave incorporating uncertainty in covariances of individual assets along with uncertainties in both variances and means in robust version of classical mean-variance model. Our special case with known moments for mean-risk models with VaR and $CVaR$ can be extended by assuming that these moments are also not known for certain or specifying the probability set differently. In this thesis, we worked only with one-period models, but for better application in practice, extensions to multi-period models could be made. There exists an excessive literature on robust approaches to portfolio selection, but in practice they are not used as often as the profound literature might suggest. As Merton states in [17, p.17], "I see this as a though engineering problem, not one of the new science".

Bibliography

[1] ALEXANDER, G. J., AND BAPTISTA, A. M. Economic implications of using a mean-VaR model for portfolio selection: a comparison with mean-variance analysis. *Journal of Economic Dynamics and Control 26* (2002), 1159–1193.

[2] AMBROŽ, L. *Měření rizika ve financích.* Ekopress, Praha, 2011.

[3] ARTZNER, P., AND ET AL. Thinking coherently: Generalised scenarios rather than var should be used when calculating regulatory capital. *Risk 10* (1997), 68–71.

[4] ARTZNER, P., AND ET AL. Coherent measures of risk. *Mathematical Finance 9* (1999), 203–228.

[5] CARVER, L. Goodbye VAR? Basel to consider other risk metrics. Availbale at `http://www.risk.net/risk-magazine/news/2154611/goodbye-var-basel-consider-risk-metrics`, February 2012. [Accessed 06.06.2014].

[6] CHOPRA, V. K., AND ZIEMBA, W. T. The effect of errors in means, variances, and covariances on optimal portfolio choice. *The journal of portfolio management 19*, 2 (1993), 6–11.

[7] DUPAČOVÁ, J., HURT, J., AND ŠTĚPÁN, J. *Stochastic Modeling in Economics and Finance.* Kluwer Academic Publisherst, Dordrech, 2002.

[8] FABOZZI, F. J., HUANG, D., AND ZHOU, G. Robust portfolios: contributions from operations research and finance. *Annals of Operations Research 176*, 1 (2010), 191–220.

[9] FED. Historical Yields of US Treasury Bonds. Availbale at `http://www.federalreserve.gov/releases/h15/data.html`. [Accessed 06.06.2014].

[10] GALICHON, A. The VAR at risk. *International Journal of Theoretical and Applied Finance 13*, 4 (2010), 503–506.

[11] GOLDFARB, D., AND IYENGAR, G. Robust portfolio selection problems. *Mathematics of Operations Research 28*, 1 (2003), 1–38.

[12] GOTOH, J., SHINOZAKI, K., AND TAKEDA, A. A Robust Portfolio Technique for Mitigating the Fragility of CVaR Minimization and generalization to coherent risk measures. *Quantitative Finance iPrint* (2013).

[13] HIRSHLEIFER, J. Investment decision under uncertainty: choice-theoretic approaches. *The Quarterly Journal of Economics* (1965), 510–536.

[14] JORION, P. *Value-at-risk: the new benchmark for managing financial risk.* McGraw-Hill, New York, 2006.

[15] MARKOWITZ, H. M. Portfolio Selection. *The Journal of Finance 7* (1952), 77–91.

[16] MARKOWITZ, H. M. *Portfolio selection: Efficient Diversification of Investments.* John Wiley & Sons, New York, 1959.

[17] MERTON, R. C. Thoughts on the future: theory and practice in investment management. *Financial Analysts Journal* (2003), 17–23.

[18] NEUMANN, J. V., AND MORGENSTERN, O. *Theory of Games and Economic Behavior.* Princeton University Press, Princeton, 1944.

[19] PFLUG, G. C. Some remarks on the value-at-risk and the conditional value-at-risk. In *Probabilistic constrained optimization*, S. Uryasev, Ed. Springer, 2000, pp. 272–281.

[20] PFLUG, G. C., AND RÖMISCH, W. *Modeling, Measuring and Managing Risk.* World Scientific Publishing Company, Singapore, 2007.

[21] R CORE TEAM. R: A Language and Environment for Statistical Computing. Available at `http://www.R-project.org/`, 2014.

[22] ROCKAFELLAR, T. R., AND URYASEV, S. Optimization of conditional value-at-risk. *Journal of Risk 2* (2000), 21–41.

[23] ROCKAFELLAR, T. R., AND URYASEV, S. Conditional Value-At-Risk for General Loss Distributions. *Journal of Banking and Finance 26* (2002), 1443–1471.

[24] RUSTEM, B., BECKER, R. G., AND MARTY, W. Robust min–max portfolio strategies for rival forecast and risk scenarios. *Journal of Economic Dynamics and Control 24*, 11 (2000), 1591–1621.

[25] SARYKALIN, S., SERRAINO, G., AND URYASEV, S. Value-at-Risk vs. Conditional Value-at-risk in Risk Management and Optimization. *Tutorials in Operations Research. INFORMS, Hanover, MD* (2008), 270–294.

[26] TURLACH, B. A., AND WEINGESSEL, A. quadprog: Functions to solve quadratic programming problems. Available at `http://CRAN.R-project.org/package=quadprog`, 2013. R package version 1.5-5.

[27] TÜTÜNCÜ, R. H., AND KOENIG, M. Robust asset allocation. *Annals of Operations Research 132*, 1-4 (2004), 157–187.

[28] ČERBÁKOVÁ, J. *Incomplete Information in Stochastic Programming Problems.* PhD thesis, MFF UK, Praha, 2008.

[29] YAHOO!FINANCE. Historical Prices. Availbale at `http://finance.yahoo.com/q`. [Accessed 06.05.2014].

[30] ZÁKUTNÁ, T. Robustní optimalizace portfolia. Master's thesis, MFF UK, Praha, 2013.

Appendix A

Descriptions of selected companies

Barnes & Noble, Inc.

It is the largest retail bookseller in the United States. The Company is also the leading retailer of content, digital media and educational products.

The Boeing Company

An American multinational corporation that designs, manufactures, and sells fixed-wing aircraft, rotorcraft, rockets and satellites. It is the second-largest aerospace & defense contractor in the world.

Credit Suisse Group AG

A Switzerland-based multinational financial services holding company. The Company operates in three business segments: Private Banking, Investment Banking and Asset Management.

Exxon Mobil Corporation

An American multinational corporation formed in 1999 that manufactures commodity petrochemicals, including aromatics, polyethylene and polypropylene plastics and a range of speciality products.

Ford Motor Company

An American multinational company that produces cars and trucks. It sells automobiles and commercial vehicles under the Ford brand and most luxury cars under the Lincoln brand.

General Electric Company

It is a diversified technology and financial services company. The products and services of the Company range from aircraft engines and household appliances to medical imaging and industrial products.

Hewlett-Packard Company

An American multinational information technology corporation. It provides products, technologies, hardware, software, solutions and services to various clients.

McDonald's Corporation

This company founded in 1940 franchises and operates McDonald's restaurants in 119 countries in the world. It is the largest chain of hamburger fast food restaurants.

PepsiCo, Inc.

An American multinational food and beverage company formed in 1965. Main activities are manufacturing, marketing and distribution of grain-based snack foods, beverages, and other products.

Pfizer Inc.

An American multinational research-based, global biopharmaceutical company founded in 1849. It is one of the world's largest pharmaceutical companies by revenues.

Wal-Mart Stores, Inc.

An American multinational retail corporation founded in 1962 that runs chains of large discount department stores and warehouse stores. It is the world's largest public corporation.

The Walt Disney Company

An American diversified multinational mass media corporation founded in 1923 as the Disney Brothers Cartoon Studio. It is the world's second largest broadcasting and cable company in terms of revenue.

Appendix B

Developments of share prices

We show the development of closing prices of shares of the selected companies reported weekly in the time period from 26.4.2004 until 27.4.2014.
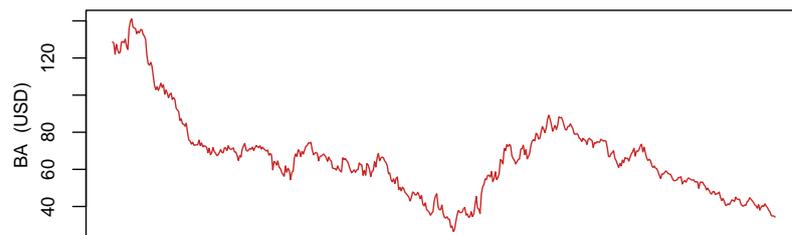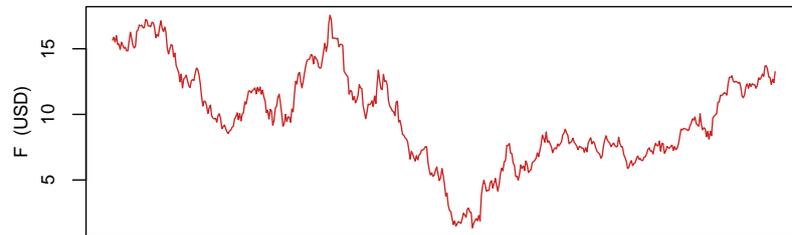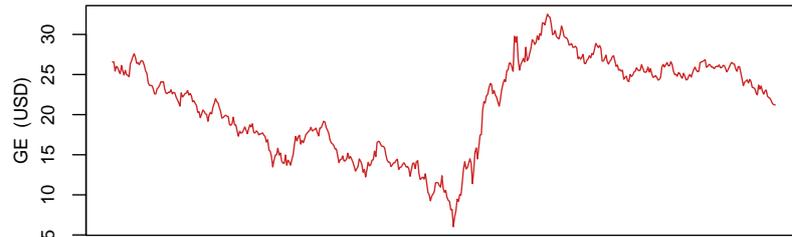

(a) Share prices of BA
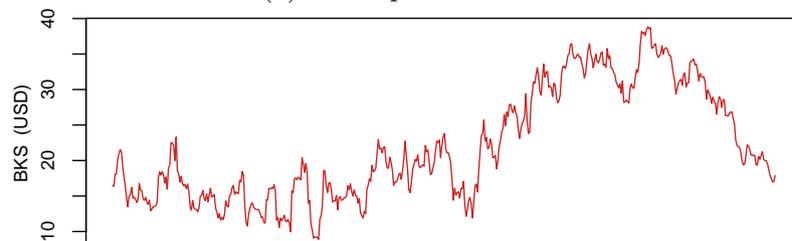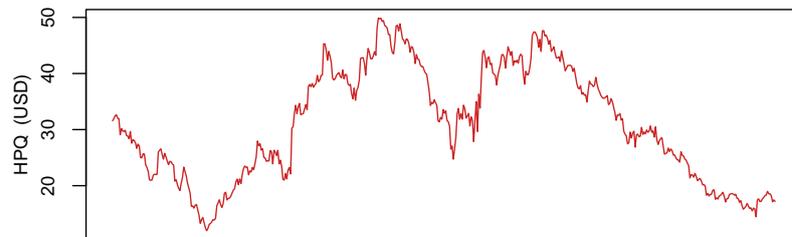

(b) Share prices of BKS
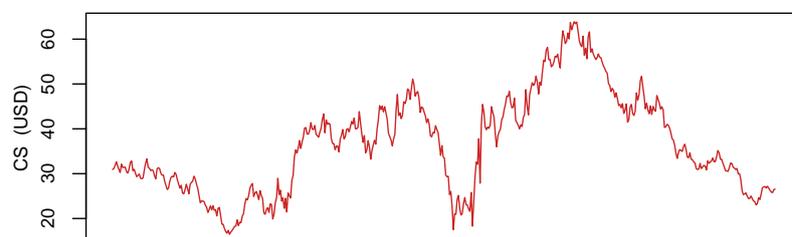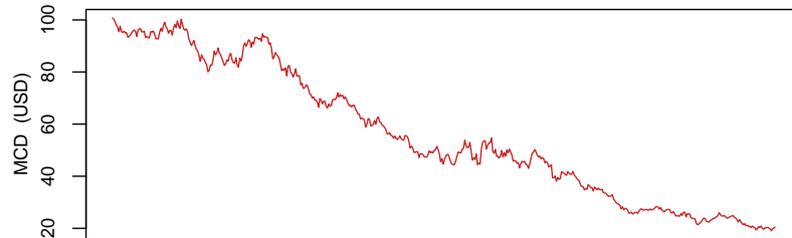

(c) Share prices of CS
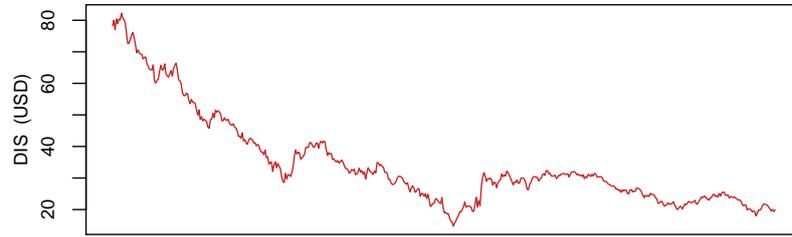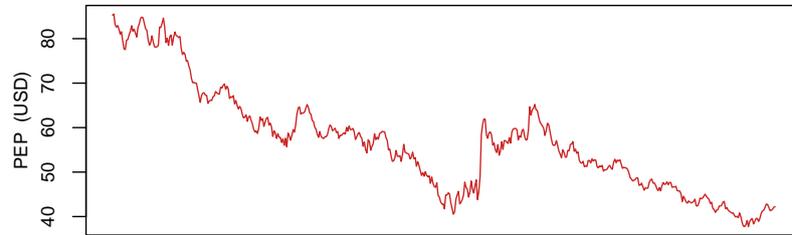
(d) Share prices of DIS

(e) Share prices of F

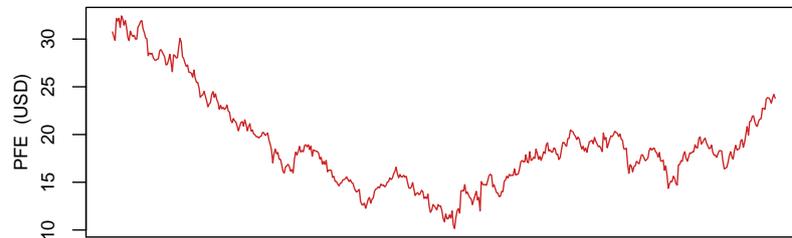
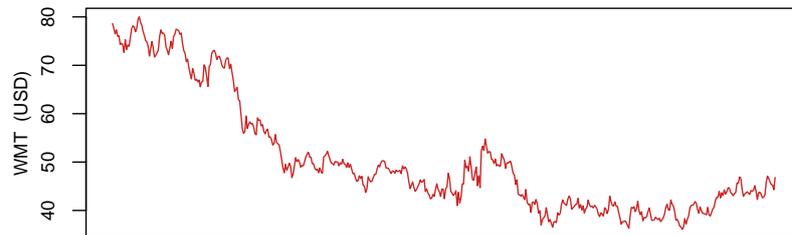(f) Share prices of GE

(g) Share prices of HPQ
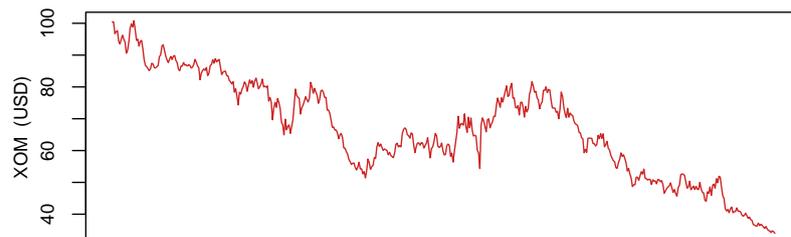

(h) Share prices of MCD

(i) Share prices of PEP

(j) Share prices of PFE

(k) Share prices of WMT

(l) Share prices of XOM

Figure B.1: Developments of share prices in the period from 26.4.2004 until 27.4.2014

Appendix C

Source code

Here we present the source code used for solving our optimization problems in R software.

```
# for solve.QP
library("quadprog")

eff.mean.variance <- function(returns, e=0.01, e.var=0.01, robust.
    var=FALSE, robust.mean=FALSE)
{       covariance <- cov(returns)
        n <- ncol(covariance)

        # add e.var to the diagonal
        if(robust.var)
        {       covariance <- covariance + diag(e.var, nrow=n, ncol
            =n)
        }

        # allocation totals to 1
        Amat <- matrix (1, nrow=n)
        bvec <- 1
        meq <- 1

        # calculate average returns
        colmn.ret <- colMeans(returns)

        # decrease average returns by e
        if(robust.mean)
        {       colmn.ret <- colmn.ret - e
        }

        # minimum return constraint
        Amat <- cbind(Amat, colmn.ret)
        bvec <- c(bvec, 0);
```

```r
        min.return.idx <- length(bvec)

        mn.ret <- mean(colmn.ret)
        mn.ret.noeps <- mean(colMeans(returns))

        # find the maximum possible return
        max.samp <- max(colmn.ret) / mn.ret
        samps <- seq(from=0.005, to=max.samp, by=0.005)

        # number of samples
        loops <- length(samps)
        loop <- 1

        # prepare the result matrices
        eff <- matrix(nrow=loops, ncol=2)
        all <- matrix(nrow=loops, ncol=n)
        colnames(eff) <- c("sd", "ret")

        # Loop through the quadratic program solver
        for (i in samps)
        {       # there is no linear component in the objective
           function
                dvec <- colMeans(returns) * 0

                # move along the efficient frontier
                bvec[min.return.idx] <- i * mn.ret

                # minimize x^T (V + epsilon * I) x
                sol <- solve.QP(covariance, dvec=dvec, Amat=Amat,
                   bvec=bvec, meq=meq)

                eff[loop,"sd"] <- sqrt(sum(sol$solution * colSums((
                   covariance * sol$solution))))
                eff[loop,"ret"] <- as.numeric(sol$solution %*%
                   colmn.ret) / mn.ret.noeps

                all[loop,] <- sol$solution

                loop <- loop+1
        }


        return(list(alloc=all, ret=eff[,"ret"], risk=eff[,"sd"]))
}
```