Portfolio Theory
Covariance and the Optimal Portfolio

d-fine

New College

University of Oxford

A thesis submitted in partial fulfillment of the requirements for the MSc in

Mathematical Finance

June 29, 2003

Abstract

The last ten years with the boom and the crash of the new economy drastically proofed the importance of considering not only the potential return but also the inherent risk of any investment in exchange traded assets.

A first mathematical approach to analyze investment opportunities in a given market considering both of these two factors where made in 1952 by Markowitz [5]. Starting from this work the single and multi-index models were derived shortly afterward (see e.g. Elton & Gruber [1]). Although those models overcame several limitations of the first model, they assume like Markowitz, that the necessary input parameters return, risk and covariance of assets respectively indices can be measured exactly.

Since the first days of the models a broad range of techniques were developed to find better estimators for statistical data, but it was only 2001 that Goldfarb and Iyengar [2] reformulated investment problems as robust problems based on uncertainty sets. Thus, instead of inserting a single data point for each input parameter, each input parameter is now represented by a set and the search for an optimal solution is considered as a worst case problem.

Furthermore, this robust formulation is equivalent to a second order cone problem and can be solved with available numerical software. In this Master Thesis all steps to calculate the minimum variance problem as an second order cone problem are presented. Then the robust minimum variance problem is calculated on actual market data of the last three years and also on simulated data. To check whether Goldfarb's and Iyengar's new approach really is an improvement, the results are compared to the classical solution of the same problem on this data.

Contents

List of Tables

Chapter 1

Introduction

In the late 1990'th, when the stock markets where booming, especially in the new economy sector, the bookshops and magazines where bursting with books and articles promising to show the way to multiply ones money easily without any effort. Parts of the so called expert tips came close to a modern search for the Holy Grail. Now about two year's after the end of the boom, many people have learned that it is not only possible to earn money by investing into assets but that there always exists some risk in loosing parts or all of the investment.

Therefore, rational investors always take into account not only the possible or expected return on an investment but also the risk that it can turn out badly. The return is usually expressed as the percentage of additional wealth one expects to earn by ones investment. For traded assets like stocks, where a history of prices is available, a straight forward and wildly used measure for risk is the historic volatility of the prices.

Within this risk-return space the search for the optimal portfolio was redefined as soon as 1952 by Markowitz [5]. He published the first mathematical approach to calculate portfolios considering both risk and return. I will introduce this famous idea in Chapter 2. Although the Markowitz approach was the breakthrough for a mathematical portfolio theory this first model has several limitations. Therefore the more advanced single and multi index models were developed which will be described in Section 2.5. Assuming that a market is driven by a limited number of factors and can hence be represented by a multi-regression model for given set of indices, these models were able to reduce the number of necessary input parameters (and computational complexity). Nevertheless, these models still treat the estimated input parameters as if they were exact. Thus single and multi-index models are still vulnerable to estimation errors in the input parameters.

One way to improve the results of the portfolio models is probably to use more advanced technologies to estimate the input data, e.g. GARCH, ARCH or ARIMA-models instead of the straight forward estimation on historical data of the input parameters. Nevertheless, this is not the aim of this thesis and therefore I will always use the standard estimators on historical data for the input parameters which will be shortly described in Chapter 3.

So instead of improving the estimation of the input parameters I will follow the recent work of Goldfarb and Iyengar [2] and except the input parameters as inexact but known to lay within certain bounds - so called uncertainty sets - with a given probability. This is also described in Section 3.4.

Working with uncertainty sets instead of "exact" values enforces a reformulate of the investment problems. So instead of searching the best portfolio with a given risk or return, a conservative worst case strategy is implemented, i.e. the portfolio with the lowest worst case risk or the highest worst case return is calculated. This robust formulation of the investment problems is done in Section 4.

The robust description of the investment problems with the derived uncertainty sets can be reformulated as a second order cone problem, i.e. they can be replaced by an optimization equation and a set of equality, inequality and cone constraints. This is demonstrated on the minimum variance problem in Chapter 4.2.

In Chapter 5 the results on some computational experiments are presented. Based on data from 06.07.1999 to 30.05.2003 of assets in the HDAX I calculated the mean variance portfolio using the classical and the robust formulation and compared the performance with respect to different aspects like the mean variance or the necessary shifts in the portfolio.

Finally, Chapter 6 completes this master thesis with a summary and an outlook on possible further studies on the subject.

Chapter 2

Basic Concepts of Portfolio Theory

When trading one always invests money in the prospect of some positive return. But with most investments one can't be sure about the return and there is a certain risk of receiving less return than hoped for or even loosing some or all of the invested money.

When pricing risky assets with a Black-Scholes model one always sets up a riskless portfolio by theoretically combining one risky asset A with another risky asset B that behaves like the opposite of A, i.e. the combined portfolio of asset A and asset B earns the riskfree rate. This investment strategy is called hedging and is widely used by banks. One of the advantages a bank has in comparison to its customers is a better access to financial markets. Therefore a bank buys an asset slightly below market price, hedges away the risk and is earns a riskless profit.

Fund managers on the other hand aim to outperform the riskless return rate for their costumers. According to the theory of efficient markets this can only be achieved by taking risk. The basic question of portfolio theory is therefore how to make the most out of a given investment when one is allowed to take a certain amount of risk. Of course there are a lot of financial and legal boundary conditions like the fund managers access to financial markets, the interests of the customer, tax regulations, trading laws etc. In this work we will focus on the mathematical part of portfolio theory, i.e., we start with the following assumptions:

- The fund manager has decided about the universe of tradable assets.
- A portfolio is only left untouched for a defined period of time and we are interested of its return after that period.
- The behaviour of the asset price S is modeled as a geometrical Brownian motion:

$$dS = \mu S dt + \sigma S dX$$

  where $d <>$ denotes the change of a variable $<>$ during the time step dt, μ is the expected return and σ the standard deviation of the return of the asset price. The stochastic process X is a Brownian motion.

- The covariance σ_{ij} between two assets S_i and S_j can be measured

Obviously the fund manager has no influence on the expected return μ or the volatilities σ of each single assets. However he can decide about how much of his funds to invest into one single asset. It is rare that an investor spends all his money on a single security. Instead, a portfolio consists usually of a combination of different assets, which is known as asset diversification. We will see later that with this strategy it is possible to reduce the risk of the portfolio without sacrificing the expected return. As most people prefer high returns and are adverse to the risk of loosing their money, diversification is the obvious thing to do.

Suppose we have decided to trade within a universe of N risky assets with prices S_i, where short selling is allowed. Then we have to decide which of the N assets we want to take into our portfolio and what proportion ϕ_i we want to spend on an asset S_i. As short selling is allowed, some of the weights ϕ_i can be negative. The set of all possible portfolios Π which can be constructed is called the opportunity set. The fund manager thus faces the task of finding the optimal portfolio for his client within this set. Markowitz was first to publish a mathematical solution to this problem in 1952 [5].

2.1 Markowitz Theory

Let there be n risky assets with prices $S_1, \ldots, S_n$ and expected returns $\mu_1, \ldots, \mu_n$. Let $V = (\sigma_{ij})$ with $i, j = 1, \ldots, n$ denote the covariance matrix of the assets. Then any portfolio Π of the opportunity set can be described by:

$$\Pi = \sum_{i=1}^{n} \phi_i S_i \qquad \text{subject to the constraint} \qquad \sum_{i=1}^{n} \phi_i = 1$$

The restriction on the ϕ_i just reflects the assumption that the entire available wealth is invested. When short selling is are allowed the weights ϕ_i can be any real number. Investor are not usually interested in the opportunity set itself but in the possible pairs of risk and return it offers. The opportunity set is therefore usually visualized by a risk-return diagram like the one in Figure (2.1). In this example the risk-return pairs of five Dax-stocks given in Table (2.1) are plotted. To visualize the risk-return pair not only of single assets but or the portfolio Π, one has to calculate the expected return μ_Π and the risk σ_Π as follows:

$$\mu_\Pi = \sum_{i=1}^{n} \phi_i \mu_i \qquad \text{and} \qquad \sigma_\Pi = \sqrt{\sum_{i,j=1}^{n} \phi_i \sigma_{ij} \phi_j}$$

The covariance matrix $V = (\sigma_{ij})$ is positive definite and therefore the set of all possible $[\sigma, \mu]$ tuples is convex. Of course not every portfolio in the opportunity is interesting for an investor. A portfolio A is said to be dominated by a portfolio B if

i	name	R	σ
1	Allianz AG	0.009	0.0078
2	BASF AG	0.0056	0.0111
3	RWE AG	0.0019	0.0084
4	Siemens AG	0.0014	0.0070
5	Volkswagen AG	0.0028	0.0077

Table 2.1: risk and return of 5 assets on 04-Okt-1996

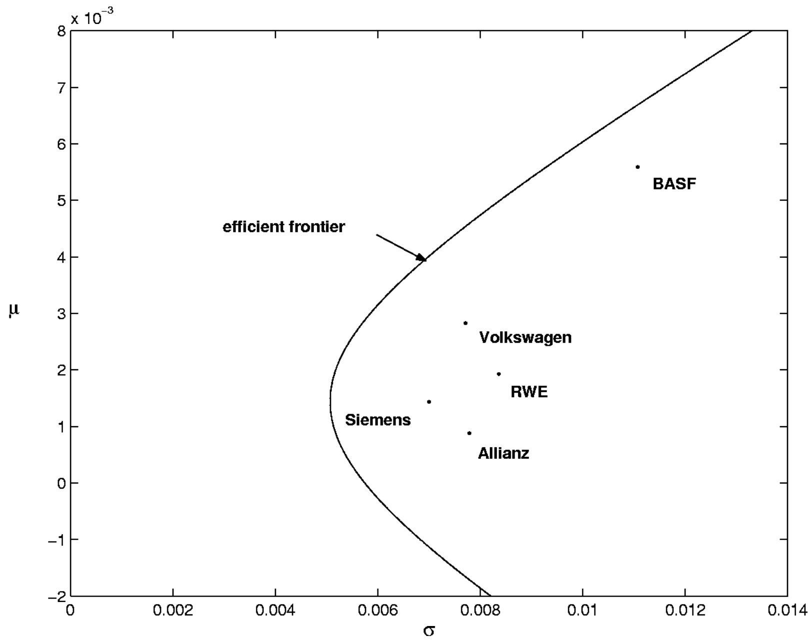


Figure 2.1: risk-return-diagram

- portfolio A is as risky as portfolio B but has a lower expected return, i.e.

$$\sigma_A = \sigma_B \text{ and } \mu_A \leq \mu_B \text{ ,or}$$

- portfolio A is expected to earn the same return as portfolio B but takes a greater risk, i.e.

$$\mu_A = \mu_B \text{ and } \sigma_A \geq \sigma_B$$

Obviously, an investor would prefer to invest in portfolios which are not dominated by others. These are called efficient portfolios. Analyzing the diagram above, we see that all efficient portfolios lie on the upper left-hand boundary of the opportunity set. The set of all efficient portfolios is called the efficient set or efficient frontier.

To find the efficient frontier Markowitz searched for any given return R_{fix} the portfolio Π of the opportunity set with the lowest risk. To simplify notation we will use the following vectors and matrices, where ^{T} denotes the transposed vector or matrix:

$$\begin{aligned}
\phi &= [\phi_1, \ldots, \phi_n]^T \\
R &= [\mu_1, \ldots, \mu_n]^T \\
V &= (\sigma_{ij}) \quad \text{with} \quad i, j = 1, \ldots, n \\
E &= [1, \ldots, 1]^T
\end{aligned}$$

So we have to solve the following constrained minimization problem:

$$\begin{array}{lll}
\text{minimize} & \sigma^2 = \phi^T V \phi & \\
\text{subject to} & R^T\phi & = \; R_{\mathit{fix}} \\
 & \sum_{i=1}^{n} \phi_i & = \; 1
\end{array}$$

The Lagrangian for this problem is:

$$\begin{aligned}
\mathcal{L}(\phi, \alpha, \beta) &= \phi^T V \phi - \alpha(R^T\phi - R_{\mathit{fix}}) - \beta(E^T\phi - 1) \\
&= \sum_{i,j=1}^{N} \phi_i \sigma_{ij} \phi_j - \alpha \sum_{i=1}^{N} R_i \phi_i + \alpha R_{\mathit{fix}} - \beta \sum_{i=1}^{N} \phi_i + \beta
\end{aligned}$$

The KKT equations of this problem are the following:

$$\frac{\partial L}{\partial \phi_i} = \sum_{j=1}^{n} \sigma_{ij}\phi_j + \phi_j \sigma_{ji} - \alpha R_i - \beta = 0, \quad (i = 1, \ldots, n) \tag{2.1}$$

$$R_{\mathit{fix}} - R^T\phi = 0 \tag{2.2}$$

$$1 - E^T\phi = 0, \tag{2.3}$$

where (2.2) and (2.3) assume feasiblity and (2.1) stationarity. The covariance matrix V is symmetric. Thus we can write (2.1) in matrix notation as follows:

$$2V\phi = \alpha R + \beta E$$

Rearranging for ϕ yields:

$$\phi = \frac{1}{2} V^{-1} R \alpha + \frac{1}{2} V^{-1} E \beta$$

Inserting this result into (2.2) and (2.3) gives:

$$\begin{aligned}
R^T\phi &= \frac{\alpha}{2} R^T V^{-1} R + \frac{\beta}{2} R^T V^{-1} E = R_{\mathit{fix}} \\
E^T\phi &= \frac{\alpha}{2} E^T V^{-1} R + \frac{\beta}{2} E^T V^{-1} E = 1
\end{aligned}$$

Let us define:

$$A = \frac{1}{2} R^T V^{-1} R; \qquad B = \frac{1}{2} R^T V^{-1} E; \qquad D = \frac{1}{2} E^T V^{-1} E$$

Using the arithmetic rules for transposed matrices

$$(MN)^T = M^T N^T \quad \text{and} \quad (M^{-1})^T = (M^T)^{-1}$$

we get:

$$R^T V^{-1} E = ((V^{-1})^T R)^T E = E^T ((V^T)^{-1} R) = E^T V^{-1} R$$

With this we receive the following matrix system for α, β:

$$\begin{pmatrix} A & B \\ B & D \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} R_{fix} \\ 1 \end{pmatrix}$$

Inversion results in:

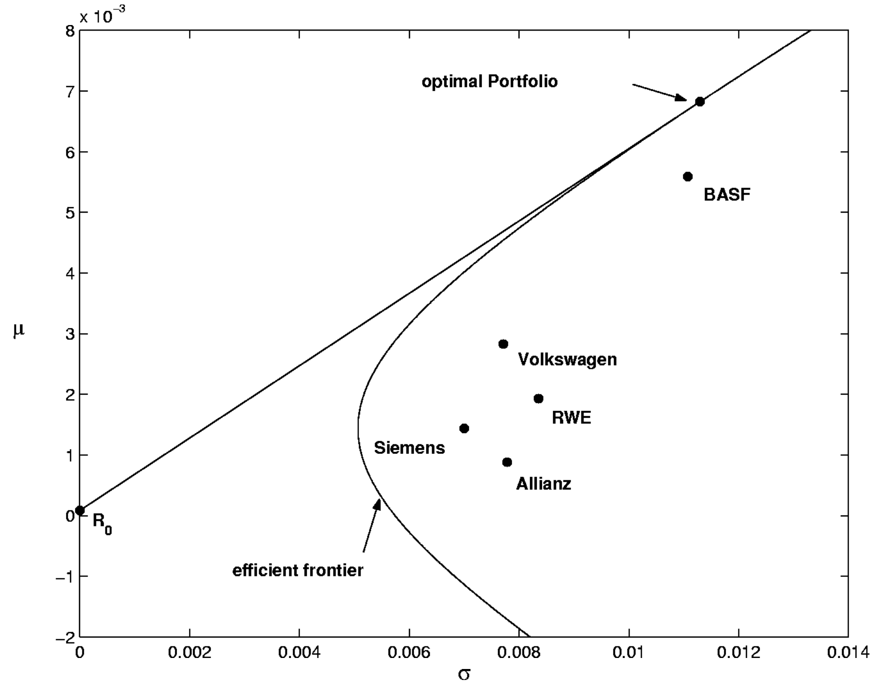
$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \frac{1}{(AD - B^2)} \begin{pmatrix} D & -B \\ -B & A \end{pmatrix} \begin{pmatrix} R_{fix} \\ 1 \end{pmatrix}$$

With this we can now calculate the efficient frontier for each value of R_{fix} :

$$\begin{aligned}
2\phi^T V \phi &= (\frac{\alpha}{2} R^T V^{-1} + \frac{\beta}{2} E^T V^{-1})(\alpha R + \beta E) = \\
&= \frac{\alpha^2}{2} R^T V^{-1} R + \frac{\alpha\beta}{2} R^T V^{-1} E + \frac{\alpha\beta}{2} E^T V^{-1} R + \frac{\beta^2}{2} E^T V^{-1} E \\
&= A\alpha^2 + 2B\alpha\beta + D\beta^2 \\
&= \frac{1}{(AD - B^2)^2} \left[A(DR_{fix} - B)^2 + 2B(DR_{fix} - B)(A - BR_{fix}) + D(A - BR_{fix})^2 \right] \\
&= \frac{1}{(AD - B^2)^2} [(AD^2 R_{fix}^2 - 2ABDR_{fix} + AB^2) + \\
&\quad +(2ABDR_{fix} - 2B^2 DR_{fix}^2 - 2AB^2 + 2B^3 R_{fix}) \\
&\quad +(A^2 D - 2ABDR_{fix} + B^2 DR_{fix}^2)] \\
&= \frac{1}{(AD - B^2)^2} [AD(DR_{fix}^2 - 2BR_{fix} + A) \\
&\quad -B^2(A + DR_{fix}^2 - 2BR_{fix}) \\
&= \frac{1}{(AD - B^2)^2} (AD - B^2)(DR_{fix}^2 - 2BR_{fix} + A) \\
&= \frac{1}{AD - B^2} (DR_{fix}^2 - 2BR_{fix} + A)
\end{aligned}$$

2.2 The maximum Sharpe ratio problem

Until now we have only allowed risky assets in our asset universe. Lets now consider a riskless asset with the riskfree rate R_0 - or at least an asset with inconsiderable risk - like a government bond of a first world country. Like we did with the risky securities we allow the riskfree asset to be bought or sold short, i.e., we assume that money can be borrowed or lent at the same riskfree rate R_0. Our portfolio opportunity set is then any linear combination of the riskless asset with any portfolio of our previous opportunity set. If we look again at our (σ, μ) diagram this is represented by the straight line through the points $(0, R_0)$ and (σ_Π, μ_Π).



When looking at the above picture we see immediately that the new efficient set of portfolios is described by the tangent to the old efficient frontier through the point $(0, R_0)$ where $\mu \geq 0$. This new efficient set is also known as the capital market line and the risky portfolio M represented by the point of contact of the tangent on the efficient frontier (σ_M, μ_M) is called the optimal portfolio. This means any efficient portfolio consists of a linear combination of the riskfree asset and the optimal portfolio M.

Thus, adding a riskless asset changes our task into finding the optimal portfolio instead of the efficient frontier. We can achieve this by finding the tangent to the efficient frontier. The tangent is the linear combination ϕ of our riskless asset with a risky portfolio of the risky opportunity set so as to maximize with the slope. The slope can be described as a function Ψ of ϕ:

$$\Psi(\phi) = \frac{R^T\phi - R_0 E^T\phi}{\sqrt{\phi^T V \phi}}$$

So, in a first step we have to solve the following constrained maximization problem:

$$\begin{array}{ll} \text{maximize} & \Psi \\ \text{subject to} & \sum_{i=1}^{n} \phi_i = 1 \end{array}$$

To simplify notation lets denote:

$$\begin{aligned} X &= (R^T - R_0 E^T)\phi \\ Y &= \phi^T V \phi \end{aligned}$$

Take partial derivatives of Ψ-function with respect to ϕ and set equal to zero:

$$\frac{\partial \Psi}{\partial \phi} = \frac{1}{Y}\left(\frac{\partial X}{\partial \phi} Y^{1/2} - X \frac{\frac{\partial Y}{\partial \phi}}{2Y^{1/2}}\right) = 0$$

This is equivalent to:

$$2\frac{\partial X}{\partial \phi} = \frac{X}{Y}\frac{\partial Y}{\partial \phi}$$

The derivatives of X and Y are:

$$\frac{\partial X}{\partial \phi} = R^T - R_0 E^T \quad \text{and} \quad \frac{\partial Y}{\partial \phi} = 2V\phi$$

Insert this into (2.2):

$$(R_i - R_0) = \frac{X}{Y} V\phi$$

Let Z denote

$$Z = \frac{X}{Y}\phi$$

We receive the system of equations:

$$VZ = (R - R_0 E)$$

Solve this to receive $Z_i, i = 1, \ldots, n$. Then calculate ϕ_i knowing that:

$$\sum Z_i = \sum \phi_i \frac{X}{Y} = \frac{X}{Y}\sum \phi_i = \frac{X}{Y}$$

so that

$$\frac{Z_i}{\sum Z_i} = \frac{\frac{X}{Y}\phi_i}{\frac{X}{Y}} = \phi_i$$

Our optimal portfolio M is then

$$M = \{\phi_0 S_0, ..., \phi_n S_n\}$$

with the expected return and risk:

$$\mu_M = \phi_M^T R \qquad \text{and} \qquad \sigma_M = \sqrt{\phi_M^T V \phi_M}$$

Note that when we derived the solution we only searched for an extrema. There is no guarantee that the solution maximizes the slope. It can as well minimize the slope and result in a portfolio with negative expected return. Thus the maximum Sharpe ratio problem does not always lead to a sensible portfolio.

2.3 Minimum variance and maximum return problem

As the maximum Sharpe ratio problem does not always have a solution, the search for the optimal portfolio is often either formulated as the minimum variance or the maximum return problem. The objective of the minimum variance problem is to find the portfolio with the minimum variance in the set of all portfolios that have a specified expected return of α. Analog the maximum return problem is the task to identify within a set of portfolios with a given risk λ the one that has the largest expected return.

2.4 Limitations of the Markowitz approach

Although the idea behind the Markowitz model is bribing as it represents a mathematical approach to measure objective investment decisions, it is not of practical value. There are several reasons for this. The most obvious is the sheer number of necessary input parameter. Reconsider that we need an estimate of the expected return μ_i and the risk σ_i of each security in the opportunity set, plus estimates of the correlation between each pair of securities. This is a total of $2n+n*(n-1)/2$ values. All of them can not be known exactly and may change over time. The quality of the parameter estimation will influence our results substantial.

Historical the first improvement of the Markowitz model where the index models which reduce the number of necessary input parameters to a manageable number as will be described in the next section.

2.5 The Single-Index Model

The single-index model is based on the observation that stocks tend to move with the market, i.e. when a market - represented by a market index - goes either up or down so do most stock prices. Therefore, it can be expected that one source of correlation between stock prices is a similar reaction to market changes. In fact, the single-index model assumes that the only relevant cause of correlation between two assets is their common dependency on market movements.

2.5.1 Definition of the Single-Index Model

The mathematical model for this assumption is a two variable linear regression equation with the return on each market stock as dependent variable and the return on a stock market index as independent variable:

$$R_i = \alpha_i + \beta_i R_m + \epsilon_i \text{ for all } i = 1, \ldots, n$$

where

R_i is the return of security S_i

R_m denotes the return on the market index

α_i is the part of security i's return which is insensitive to market changes

β_i measures the effect of changes in the market return R_m on the securities return

ϵ_i represents the uncertainty in the securities return R_i

The linear regression model splits the return on a stock into a part dependent on the market and a part which is only dependent on the security itself. Both the market return R_m and the error term ϵ_i are random variables and as such have a probability distribution, a mean and a standard deviation σ_m, σ_{ϵ_i}. Regression analyis estimates the coefficients α_i, β_i, such that ϵ_i has an expected value of zero and will be uncorrelated to R_m at least for the historical time period that was used for the regression, i.e.

$$E[\epsilon_i] = 0, \qquad \text{cov}\ (\epsilon_i, R_m) = E[(\epsilon_i - 0)(R_m - \bar{R}_m)] = 0.$$

Until now all mentioned properties of the single-index model hold either by definition or by construction. The key characteristic of the single index model is the assumption that the movement of the market is the only reason for the correlation between assets. This implies that the error terms of any two different stock prices S_i, S_j are independent from each other, i.e.,

$$E[\epsilon_i \epsilon_j] = 0$$

Any effects like industry dependencies that could cause correlation between assets are neglected. However, the assumption that all other possible sources of correlation are inconsiderable small cannot be forced to be true neither by definition or construction. Thus the quality of the single-index model is dependent on how much this simplification differs from reality. So altogether the single-index model has the following properties:

Single-Index Model:

Basic Equation:	$R_i = \alpha_i + \beta_i R_m + \epsilon_i$
By Construction:	$E[\epsilon_i] = 0$
By Definiton:	Variance of ϵ_i is denoted by $\sigma^2_{\epsilon_i}$ Variance of R_m is denoted by σ^2_m
By Assumption:	$E[\epsilon_i(R_m - \bar{R}_m)] = 0$, i.e., index independent of asset specific part of return $E[\epsilon_i \epsilon_j] = 0$, securities only correlated by common response to market

In the next section we will see how this model simplifies the estimation of the input parameters.

2.5.2 Mean return, variance and correlation

Lemma :
According to the single-index model the parameters of each security can be calculated by:

$$\begin{array}{llcl} \text{mean return} & \bar{R}_i & = & \alpha_i + \beta_i \bar{R}_m \\ \text{variance} & \sigma_i^2 & = & \beta_i^2 \sigma_m^2 + \sigma_{\epsilon_i}^2 \end{array}$$

The covariance of returns between securities i and j is given by $\sigma_{ij} = \beta_i \beta_j \sigma_m^2$

Proof:

Since α_i,β_i are constants and $E[\epsilon_i] = 0$ by construction, we find for the expected return on a security

$$E[R_i] = E[\alpha_i + \beta_i R_m + \epsilon_i] = E[\alpha_i] + E[\beta_i R_m] + E[\epsilon_i] = \alpha_i + \beta_i \bar{R}_m$$

The variance of the return on any security is

$$\begin{aligned}
\sigma_i^2 = E[R_i - \bar{R}_i]^2 &= E[(\alpha_i + \beta_i R_m + \epsilon_i) - (\alpha_i + \beta_i \bar{R}_m)^2] \\
&= E[(\beta_i(R_m - \bar{R}_m) + \epsilon_i)^2] \\
&= \beta_i^2 E[R_m - \bar{R}_m)^2] + 2\beta_i E[(R_m - \bar{R}_m)\epsilon_i] + E[(\epsilon_i)^2]
\end{aligned}$$

By assumption $E[(R_m - \bar{R}_m)\epsilon_i] = 0$. Thus,

$$\sigma_i^2 = \beta_i^2 E[(R_m - \bar{R}_m)^2] + E[(\epsilon_i)^2] = \beta_i^2 \sigma_m^2 + \sigma_{\epsilon_i}^2$$

Substituting the results for the mean return into the definition of the covariance between security S_i and S_j yields

$$\begin{aligned}
\sigma_{ij} &= E[(R_i - \bar{R}_i)(R_j - \bar{R}_j)] \\
&= E[(\alpha_i + \beta_i R_m + \epsilon_i - (\alpha_i + \beta_i \bar{R}_m))(\alpha_j + \beta_j R_m + \epsilon_j - (\alpha_j + \beta_j \bar{R}_m))] \\
&= E[(\beta_i(R_m - \bar{R}_m) + \epsilon_i)(\beta_j(R_m - \bar{R}_m) + \epsilon_j)] \\
&= \beta_i \beta_j E[(R_m - \bar{R}_m)^2] + \beta_j E[\epsilon_i(R_m - \bar{R}_m)] + \beta_i E[\epsilon_j(R_m - \bar{R}_m)] + E[\epsilon_i \epsilon_j]
\end{aligned}$$

Since the last three terms are zero we end up with $\sigma_{ij} = \beta_i \beta_j \sigma_m^2$ □

Thus, in the single-index model the formulas for the expected return and the variance of a security S_i contain an asset specific and a market-related part while the covariance of two securities S_i and S_j depends only on the market-dependencies β_i, β_j of the two securities and the market-risk σ_m^2 itself.

With these results, the expected return of a portfolio P can be calculated by:

$$\bar{R}_P = \sum_{i=1}^{N} \lambda_i \bar{R}_i = \sum_{i=1}^{N} \lambda_i \alpha_i + \sum_{i=1}^{N} \lambda_i \beta_i \bar{R}_m$$

and the variance of a portfolio by

$$\sigma_P^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j \sigma_{ij} = \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j \beta_i \beta_j \sigma_m^2 + \sum_{i=1}^{N} \lambda_i^2 \sigma_{ei}^2$$

To use these formulas we need estimates of $\alpha_i, \beta_i, \sigma_{\epsilon_i}^2$ for each stock, plus an additional estimate of the expected return $\bar{R}_m$ and variance σ_m^2 for the description of the market. Note that there is no need for an estimate of the correlation σ_{ij} of the assets. Thus, the single-index model reduces the number of required estimates to $3N+2$ instead of the $2N+N(N-1)/2$ necessary for the Markowitz approach. Nevertheless, it is not to be expected that asset prices are only driven by one factor. Therefore, multi-index models where derived.

2.6 The Multi-Index Model

Multi-Index models are based on the same idea as the single-index model. Only the assumption, that all common movement can be explained sufficiently already by one factor is dropped. Instead it is assumed that a number of - say m - factors together explain all of the covariance of the market prices.

2.6.1 Definition of the Multi-Index Model

Thus the regression of the multi-index model is given by:

$$R_i = \alpha_i + \beta_{i1} I_1 + \cdots + \beta_{im} I_M + \epsilon_i \text{ for all } i = 1, \ldots, n$$

where I_j is the return of the j's factor $j = 1, \ldots, m$ and β_{ij} denotes the influence of the j's factor on the return R_i of asset i. The other assumptions stay the same. Thus:

Multi-Index Model:

Basic Equation:	$R_i = \alpha_i + \beta_{i1} I_1 + \cdots + \beta_{im} I_M + \epsilon_i$ for all $i = 1, \ldots, n$
By Construction:	$E[\epsilon_i] = 0$
By Definiton:	Variance of ϵ_i is denoted by $\sigma^2_{\epsilon_i}$ Variance of factors I_j is denoted by $\sigma^2_{I_j}$ Covariance matrix of the factors is $\boldsymbol{F}$
By Assumption:	$E[\epsilon_i(I_j - \bar{I}_j)] = 0$ for all j=1,...,m $E[\epsilon_i \epsilon_j] = 0$

2.6.2 Mean return, variance and correlation

The expected return, variance and the correlation are represented by the following formulas in matrix notation:

mean return	$\bar{\boldsymbol{R}}$	$=$	$\boldsymbol{\alpha} + \boldsymbol{V}\bar{\boldsymbol{I}}$
covariance	cov $(\boldsymbol{R})$	$=$	$\boldsymbol{VFV} + \boldsymbol{\sigma_\epsilon}$

The proof for this is based on the same arguments than the one of the single index model. Thus the expected return of a portfolio P is:

$$\bar{R}_P = \boldsymbol{\phi}\bar{R}_i = \boldsymbol{\phi^T\alpha} + (\boldsymbol{VI})^{\boldsymbol{T}}\boldsymbol{\phi}$$

and the variance

$$\sigma_P^2 = \boldsymbol{\phi^T V^T F V \phi} + \boldsymbol{\sigma_\epsilon^T \phi}$$

The number of input parameters of the multi-index model increases by $2m + m * (m-1)/2$ additional estimates of the return, variance and the covariance of the factors. Nevertheless it is still significantly lower than the number of parameters of the Markowitz approach as the number of factors m in a multi-index model should be significantly lower than the number of assets n.

Chapter 3

Parameter Estimation and Uncertainty Sets

In the last two chapters the markowitz μ-σ view of portfolio theory and the index models where introduced. All of these models need several estimates. Therefore this chapter includes some economics theory to specify all parameters needed. However, the only observable parameters are the historical market prices S_i of the stock i or the market indices S_{I_j}. From those the historic returns R_i for each stock and I_j of the indices can be calculated. So the most straight forward method is to assume all parameters to be constant over time and to estimate future returns and volatilities from past data. The historical returns are then the input data for a regression analysis to calculate the regression parameters α and β. In a last step, the volatility of the error term $\sigma^2_{\epsilon_i}$ can be estimated.

3.1 Returns R_i, R_m and market volatility σ^2_m

The return of an investment is usually defined as the extra amount earned on each unit originally invested. Let S_{t_i} denote the price of an asset at time t_i and $D_{[t_i,t_{i+1}]}$ the sum of all cashflows of the asset during the time period $[t_i, t_{i+1}]$, then the return of the asset S from time t_i to t_{i+1} is given by:

$$R_{[t_i,t_{i+1}]} = \frac{S_{t_{i+1}} + D_{[t_{i+1},t_i]} - S_{t_i}}{S_{t_i}}$$

From statistics we know that an unbiased estimate for the mean or expected return of an asset is then given by:

$$\bar{R} = \frac{1}{k}\sum_{j=1}^{k} R_j$$

for a sample of k values R_j. In the finance literature two estimators of the variance are used. One is the standard unbiased estimator of σ^2:

$$\hat{\sigma}^2 = \frac{1}{K-1}\sum_{j=1}^{K}(R_j - \bar{R})^2$$

But often it is also assumed that the returns R of assets are normally distributed with constant mean μ and variance σ^2. Then instead of the above standard estimator for σ^2 the maximum likelihood estimator is used. The idea behind the maximum likelihood method is to choose the value σ^2 in such a way that it maximizes the probability of the observed data to occur. Suppose now that we have observed the K values $R_1, R_2, \ldots, R_m$ of R. Then the probability density that the sample data u_i with $i = 1, \ldots, m$ occurred in exactly this order is:

$$p(R_1, R_2, \ldots, R_K) = \prod_{j=1}^{K} \left[\frac{1}{\sqrt{2\pi\sigma_R^2}} \exp\left(\frac{(\mu_R - R_i)^2}{2\sigma_R^2} \right) \right]$$

The maximum likelihood estimator $\hat{\sigma}^2$ is now the value that maximizes this probability density function. Instead of maximizing the above expression itself we are looking for the maximum of its logarithm, which is equivalent. This leads us to:

$$\ln \left(\prod_{j=1}^{K} \left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{(\mu_R - R_j)^2}{2\sigma_R^2} \right) \right] \right) = -\frac{K}{2} \ln(2\pi) - \frac{K}{2} \ln(\sigma_R^2) - \sum_{j=1}^{K} \frac{(R_j - \mu_R)^2}{2\sigma_R^2}$$

To find the maximum we've got to differentiate with respect to σ^2 and set the result equal to zero:

$$-\frac{K}{2\hat{\sigma_R^2}} + \sum_{j=1}^{R} \frac{(R_j - \mu_R)^2}{2\hat{\sigma_R^2}^2} = 0 \Longleftrightarrow \hat{\sigma_R^2} = \frac{1}{m} \sum_{i=1}^{m} (R_j - \mu_R)^2$$

3.2 Estimating historical Betas

In the single-index model the return on a stock is represented by

$$R_i = \alpha_i + \beta_i R_m + \epsilon_i$$

This equation is assumed to be valid at each point in time, even if the values of the parameters α_i, β_i, or $\sigma_{\epsilon_i}^2$ change with time. When using the observed historical returns on securities and the market as input, one usually also assumes that α_i, β_i, and σ_{ei}^2 are constant, i.e., the same linear relationship is expected to hold at each moment in time. If the variance of the error term σ_{ei}^2 were equal to zero, then α_i and β_i could be calculated from just two observations. However, the observed values for R_i and R_m are usually not hitting the line specified by the above Equation but will lie spread around it. The larger $\sigma_{\epsilon_i}^2$, the wider the observed values are scattered and the more difficult it is to find the unobservable relationship. Usually, linear regression analysis as specified in the following theorem is used for this task.

Gauss-Markov Theorem:

Suppose, a two-variable linear regression model

$$Y = a + bX + \epsilon$$

for the dependent variable X and the independent variable Y has the following properties:

1. The X's are non stochastic variables whose values are fixed.

2. $E[\epsilon] = 0$, i.e. the expected value of the error term is zero.

3. The variance $E(\epsilon^2) = \sigma_\epsilon^2$ of the error term ϵ is constant.

4. The random variables ϵ_i are statistically independent. Thus, $E[\epsilon_i \epsilon_j] = 0$ for all $i \neq j$.

Then the ordinary least-square estimators $\hat{a}$ and $\hat{b}$ of a and b for a total number of k observations X_i, Y_i are given by:

$$\hat{b} = \frac{1/K \sum_{j=1}^{K} X_j Y_j - \bar{X}\bar{Y}}{1/K \sum_{j=1}^{K} X_j^2 - (\bar{X})^2}$$

and

$$\hat{a} = \bar{Y} - \hat{b}\bar{X}$$

The estimators $\hat{a}, \hat{b}$ are the most efficient linear unbiased estimators, i.e. have the smallest variance of all linear unbiased estimators.

For multi-index models the parameters are estimated analogous by a multivariate linear regression.

3.3 Variance of the error term σ_{ϵ_i}

Suppose we have done a regression analysis for a sample of k values. Let $\hat{\epsilon}_i = Y_i - \hat{Y}_i$ be the regression residual. Then the residual variance

$$s = \hat{\sigma}^2 = \frac{\sum \hat{\epsilon}_i^2}{N-2} = \frac{\sum (Y_i - \hat{\alpha} - \hat{\beta} X_i)^2}{N-2}$$

is an unbiased, consistent estimator of the error variance. s is called the standard error of the regression.

3.4 Uncertainty sets

Until now we have always treated the input parameters as if it is possible to estimate them with inconsiderable error. In reality this is not the case. For one thing the parameters do change with time and the sample sets are relative small. In this section I will introduce a way to find uncertainty sets for all parameters of a multi-index model. The argumentation follows mainly section 5 of Goldberg and Iyengar [2].

The underlying idea is to calculate the least squares estimates of the constants $\boldsymbol{\mu}, \boldsymbol{V}$ of the market model:

$$\boldsymbol{r} = \boldsymbol{\mu} + \boldsymbol{V}^{\boldsymbol{T}} \boldsymbol{f} + \boldsymbol{\epsilon}$$

by linear regression and to use the confidence levels of this computation to define the borders of the uncertainty set of the parameters. Assume that sample data of asset and factor returns $\boldsymbol{r}, \boldsymbol{f}$ are available for k time periods and that $\sigma^2_{\epsilon_i}$ is constant. Let

$$\begin{array}{rcll} \boldsymbol{S} & = & [\boldsymbol{r^1}, \ldots, \boldsymbol{r^k}]^T & \text{denote the } k \text{ asset return vectors} \\ \boldsymbol{y_i} & = & [r_i^1, \ldots, r_i^k]^T & \text{denote the } k \text{ returns of a single asset } i \\ \boldsymbol{B} & = & [\boldsymbol{f^1}, \ldots, \boldsymbol{f^k}]^T & \text{denote the } k \text{ factor return vectors} \end{array}$$

then for a single asset the market model reads:

$$\begin{bmatrix} r_i^1 \\ \vdots \\ r_i^k \end{bmatrix} = \begin{bmatrix} \mu_i \\ \vdots \\ \mu_i \end{bmatrix} + \begin{bmatrix} f_1^1 & \ldots & f_m^1 \\ \vdots & \vdots & \vdots \\ f_1^k & \ldots & f_m^k \end{bmatrix} \begin{bmatrix} V_{1i} \\ \vdots \\ V_{1m} \end{bmatrix} + \begin{bmatrix} \epsilon_i^1 \\ \vdots \\ \epsilon_i^k \end{bmatrix}$$

To simplify notation we define:

$$\underbrace{\begin{bmatrix} r_i^1 \\ \\ \vdots \\ \\ r_i^k \end{bmatrix}}_{\boldsymbol{y_i}} = \underbrace{\begin{bmatrix} 1 & f_1^1 & \ldots & f_m^1 \\ \\ \vdots & \vdots & \vdots & \vdots \\ \\ 1 & f_1^k & \ldots & f_m^k \end{bmatrix}}_{\boldsymbol{A}} \underbrace{\begin{bmatrix} \mu_i \\ V_{1i} \\ \vdots \\ \\ V_{1m} \end{bmatrix}}_{\boldsymbol{x_i}} + \underbrace{\begin{bmatrix} \epsilon_i^1 \\ \\ \vdots \\ \\ \epsilon_i^k \end{bmatrix}}_{\boldsymbol{\epsilon_i}}$$

The aim of linear regression is to find the minimum least square estimator of $\boldsymbol{x_i}$ (compare Pindyck & Rubinfeld [6] p.107 ff.), i.e. the vector $\boldsymbol{\hat{x}_i}$ that minimizes

$$\sum_{j=1}^{k} \hat{\epsilon}_i^j = \boldsymbol{\hat{\epsilon}_i}^{\boldsymbol{T}} \boldsymbol{\hat{\epsilon}_i}$$

where

$$\boldsymbol{\hat{\epsilon}_i} = \boldsymbol{y_i} - \boldsymbol{\hat{y}_i} \quad \text{and} \quad \boldsymbol{\hat{y}_i} = \boldsymbol{A}\boldsymbol{\hat{x}_i}$$

Substitution yields:

$$\boldsymbol{\hat{\epsilon}_i^T} \boldsymbol{\hat{\epsilon}_i} = (\boldsymbol{y_i} - \boldsymbol{A}\boldsymbol{\hat{x}_i})^T (\boldsymbol{y_i} - \boldsymbol{A}\boldsymbol{\hat{x}_i}) = \boldsymbol{y_i^T} \boldsymbol{y_i} - 2\boldsymbol{\hat{x}_i^T} \boldsymbol{A^T} \boldsymbol{y_i} + \boldsymbol{\hat{x}_i^T} \boldsymbol{A^T} \boldsymbol{A} \boldsymbol{\hat{x}_i}$$

So to find the minimum, we have got differentiate the above equation with respect to $\boldsymbol{\hat{x}_i}$:

$$\frac{\partial \boldsymbol{\hat{\epsilon}_i^T} \boldsymbol{\hat{\epsilon}_i}}{\boldsymbol{\hat{x}_i}} = -2\boldsymbol{A^T} \boldsymbol{y_i} + 2\boldsymbol{A^T} \boldsymbol{A} \boldsymbol{\hat{x}_i} = 0$$

Thus:

$$\boldsymbol{\hat{x}_i} = (\boldsymbol{A^T} \boldsymbol{A})^{\boldsymbol{-1}} \boldsymbol{A^T} \boldsymbol{y_i}$$

Lemma 1

$$\boldsymbol{\hat{x}_i} \sim N(\boldsymbol{x_i}, \sigma^2_{\epsilon_i} (\boldsymbol{A^T} \boldsymbol{A})^{-1})$$

Proof:

$$\boldsymbol{\hat{x}_i} - \boldsymbol{x_i} = (\boldsymbol{A^T} \boldsymbol{A})^{\boldsymbol{-1}} \boldsymbol{A^T} \boldsymbol{y_i} - \boldsymbol{x_i} = (\boldsymbol{A^T} \boldsymbol{A})^{\boldsymbol{-1}} \boldsymbol{A^T} (\boldsymbol{A}\boldsymbol{x_i} + \boldsymbol{\epsilon_i}) - \boldsymbol{x_i} = (\boldsymbol{A^T} \boldsymbol{A})^{\boldsymbol{-1}} \boldsymbol{A^T} \boldsymbol{\epsilon_i}$$

$$E[\hat{\boldsymbol{x}}_i] - E[\boldsymbol{x}_i] = E[\hat{\boldsymbol{x}}_{\boldsymbol{i}} - \boldsymbol{x}_{\boldsymbol{i}}] = E[(\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A})^{\boldsymbol{-1}}\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{\epsilon}_{\boldsymbol{i}}] = 0$$

$$\begin{aligned} Var[\hat{\boldsymbol{x}}_{\boldsymbol{i}}] &= E[(\hat{\boldsymbol{x}}_{\boldsymbol{i}} - \boldsymbol{x}_{\boldsymbol{i}})^T(\hat{\boldsymbol{x}}_{\boldsymbol{i}} - \boldsymbol{x}_{\boldsymbol{i}})] = E[((\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{\epsilon}_{\boldsymbol{i}})^T(\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{\epsilon}_{\boldsymbol{i}}] \\ &= E[\boldsymbol{\epsilon}_{\boldsymbol{i}}{}^T\boldsymbol{A}((\boldsymbol{A}^T\boldsymbol{A})^{-1})^T(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{\epsilon}_{\boldsymbol{i}}] = E[\boldsymbol{\epsilon}_{\boldsymbol{i}}{}^T\boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{A})^{-1}(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{\epsilon}_{\boldsymbol{i}}] \end{aligned}$$

$$= \sigma_{\epsilon_i}^2(\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A})^{\boldsymbol{-1}}$$

To determine a significance level for the variance we need its probability distribution. This can be done with the help of the following lemma (Pindyck and Rubinfeld [6] p.35)

Lemma 2 *The sum of the squares of N independently distributed normal random variables with mean 0 and variance 1 is distributed as chi square with N degrees of freedom.*

We have $m+1$ independently distributed normal random variables with mean 0 and variance $\sigma_{\epsilon_i}^2(\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A})^{\boldsymbol{-1}}$. Thus we need to transform the variance to 1 to use lemma 2:

Lemma 3

$$\frac{1}{\sigma_{\epsilon_i}}A(\hat{\boldsymbol{x}}_{\boldsymbol{i}} - x_i) \sim N(0,1)$$

Proof:

$$\begin{aligned} E[\frac{1}{\sigma_{\epsilon_i}}A(\hat{\boldsymbol{x}}_{\boldsymbol{i}} - \boldsymbol{x}_{\boldsymbol{i}})] &= \frac{1}{\sigma_{\epsilon_i}}AE[(\hat{\boldsymbol{x}}_{\boldsymbol{i}} - x_i)] = 0 \\ Var[\frac{1}{\sigma_{\epsilon_i}}A(\hat{\boldsymbol{x}}_{\boldsymbol{i}} - \boldsymbol{x}_{\boldsymbol{i}})] &= E[(\hat{\boldsymbol{x}}_{\boldsymbol{i}} - \boldsymbol{x}_{\boldsymbol{i}})^T A^T \frac{1}{\sigma_{\epsilon_i}}\frac{1}{\sigma_{\epsilon_i}}A(\hat{\boldsymbol{x}}_{\boldsymbol{i}} - \boldsymbol{x}_{\boldsymbol{i}})] \\ &= \frac{1}{\sigma_{\epsilon_i}}^2 A^T AE[(\hat{\boldsymbol{x}}_{\boldsymbol{i}} - \boldsymbol{x}_{\boldsymbol{i}})^T(\hat{\boldsymbol{x}}_{\boldsymbol{i}} - \boldsymbol{x}_{\boldsymbol{i}})] \\ &= \frac{1}{\sigma_{\epsilon_i}^2}A^T A\sigma_{\epsilon_i}^2(\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A})^{\boldsymbol{-1}} = 1 \end{aligned}$$

So with lemma 2 we have that the square of $\frac{1}{\sigma_{\epsilon_i}}A(\hat{\boldsymbol{x}}_{\boldsymbol{i}} - \boldsymbol{x}_{\boldsymbol{i}})$ is distributes as chi square with $m+1$ degrees of freedom:

$$\chi = \frac{1}{\sigma_{\epsilon_i}^2}(\hat{\boldsymbol{x}}_{\boldsymbol{i}} - \boldsymbol{x}_{\boldsymbol{i}})^T(\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A})(\hat{\boldsymbol{x}}_{\boldsymbol{i}} - \boldsymbol{x}_{\boldsymbol{i}}) \sim \chi_{m+1}^2$$

For practical computation we need to replace σ_{ϵ_i} by an unbiased estimator. The standard way in regression theory is the following estimator:

$$s_i^2 = \frac{||\boldsymbol{y}_{\boldsymbol{i}} - \boldsymbol{Y}\hat{\boldsymbol{x}}_{\boldsymbol{i}}||}{k-m-1} \tag{3.1}$$

Consider another statistical lemma (see Pindyck & Rubinfeld [6] p. 37):

Lemma 4 *If X and Z are independent and distributed as chi square with N_1 and N_2 degrees of freedom, respectively, then $(X/N_1)/(Z/N_2)$ is distributed according to an F distribution with N_1 and N_2 degrees of freedom.*

Thus with $X = \frac{1}{\sigma^2_{\epsilon_i}}(\hat{\boldsymbol{x}}_{\boldsymbol{i}} - \boldsymbol{x}_{\boldsymbol{i}})^T(\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A})(\hat{\boldsymbol{x}}_{\boldsymbol{i}} - \boldsymbol{x}_{\boldsymbol{i}})$ and $Z = s_i^2/\sigma^2_{\epsilon_i}$ and $N_1 = m+1$ and $N_2 = k-m-1$ we have got a new random variable $\mathcal{Y}$:

$$\mathcal{Y} = \frac{1}{(m+1)s_i^2}(\hat{\boldsymbol{x}}_{\boldsymbol{i}} - \boldsymbol{x}_{\boldsymbol{i}})^T(\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A})(\hat{\boldsymbol{x}}_{\boldsymbol{i}} - \boldsymbol{x}_{\boldsymbol{i}})$$

that is $\mathcal{F}$-distributed with a numerator of $m+1$ degrees of freedom and a denominator of $k-m-1$ degrees of freedom. To simplify notation, we will use the following variables

$\mathcal{F}_j$: cumulative $\mathcal{F}$-distribution with j degrees in the numerator and $k-m-1$ degrees of freedom in the denominator

$c_j(\omega)$: ω-critical value of $\mathcal{F}_j$ such that $\mathcal{F}_j(c_j(\omega)) = \omega$

We found that the probability for $\mathcal{Y} \le c_{m+1}(\omega)$ is ω and thus:

$$P((\bar{\boldsymbol{x}}_{\boldsymbol{i}} - \boldsymbol{x})^T(\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A})(\bar{\boldsymbol{x}}_{\boldsymbol{i}} - \boldsymbol{x}_{\boldsymbol{i}}) \le (m+1)c_{m+1}(\omega)\sigma^2_{\epsilon_i}) = \omega$$

So

$$S_i(\omega) = \{(\bar{\boldsymbol{x}}_{\boldsymbol{i}} - \boldsymbol{x})^T(\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A})(\bar{\boldsymbol{x}}_{\boldsymbol{i}} - \boldsymbol{x}_{\boldsymbol{i}}) \le (m+1)c_{m+1}(\omega)\sigma^2_{\epsilon_i}\}$$

is an uncertainty set for the vector $\boldsymbol{x}_{\boldsymbol{i}}$ corresponding to the probability ω. As we assume that the error terms σ_{ϵ_i} are independent, we can combine the confidence sets of the n-assets to a ω^n-confidence set for the whole asset opportunity set by:

$$S(\omega) = S_1(\omega) \times \cdots \times S_n(\omega)$$

For portfolio analysis we prefer separate uncertainty sets for μ and $\boldsymbol{V}$. Therefore we need to project the above set along the vector μ and the vector $\boldsymbol{V}_{\boldsymbol{i}}$. This can be achieved for any nonsingular matrix equation of the form $\boldsymbol{M\xi} = \boldsymbol{\zeta}$ with the use of a projection matrix $\boldsymbol{P}$ as follows:

$$\boldsymbol{M\xi} = \boldsymbol{\zeta} \quad \Leftrightarrow \quad \boldsymbol{\xi} = \boldsymbol{M}^{-1}\boldsymbol{\zeta} \quad \Leftrightarrow \quad \boldsymbol{P\xi} = \boldsymbol{PM}^{-1}\boldsymbol{\zeta} \quad \Leftrightarrow \quad (\boldsymbol{PM}^{-1})^{-1}\boldsymbol{P\xi} = \boldsymbol{\zeta}$$

This implies the following quadratic function:

$$[(\boldsymbol{PM}^{-1})^{-1}\boldsymbol{P\xi}]^T[(\boldsymbol{PM}^{-1})^{-1}\boldsymbol{P\xi}] \quad \Leftrightarrow \quad [(\boldsymbol{P\xi})^T((\boldsymbol{PM}^{-1})^T)^{-1}][(\boldsymbol{PM}^{-1})^{-1}\boldsymbol{P\xi}]$$

$$\Leftrightarrow \quad [(\boldsymbol{P\xi})^T((\boldsymbol{PM}^{-1})(\boldsymbol{PM}^{-1})^T)^{-1}\boldsymbol{P\xi}] \quad \Leftrightarrow \quad [(\boldsymbol{P\xi})^T(\boldsymbol{P}^T(\boldsymbol{M}^T\boldsymbol{M})^{-1}\boldsymbol{P})^{-1}\boldsymbol{P\xi}]$$

Reconsider the entries of the matrices in the quadratic function above

$$\begin{bmatrix} \hat{\boldsymbol{\mu}}_{\boldsymbol{0}} - \boldsymbol{\mu}_{\boldsymbol{0}} \\ \hat{\boldsymbol{V}}_{\boldsymbol{1i}} - \boldsymbol{V}_{\boldsymbol{1i}} \\ \vdots \\ \hat{\boldsymbol{V}}_{\boldsymbol{mi}} - \boldsymbol{V}_{\boldsymbol{mi}} \end{bmatrix} \underbrace{\begin{bmatrix} 1 & 1 & \dots & 1 \\ f_1^1 & f_1^2 & \dots & f_1^k \\ \vdots & \vdots & \vdots & \vdots \\ f_m^1 & f_m^2 & \vdots & f_m^k \end{bmatrix}}_{\boldsymbol{A}^{\boldsymbol{T}}} \underbrace{\begin{bmatrix} 1 & f_1^1 & \dots & f_m^1 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 1 & f_1^k & \dots & f_m^k \end{bmatrix}}_{\boldsymbol{A}} \begin{bmatrix} \hat{\boldsymbol{\mu}}_{\boldsymbol{0}} - \boldsymbol{\mu}_{\boldsymbol{0}} \\ \hat{\boldsymbol{V}}_{\boldsymbol{1i}} - \boldsymbol{V}_{\boldsymbol{1i}} \\ \vdots \\ \hat{\boldsymbol{V}}_{\boldsymbol{mi}} - \boldsymbol{V}_{\boldsymbol{mi}} \end{bmatrix}$$

Let $\boldsymbol{e_j}$ denote the j-s identity-vector. The projection matrix for μ is then $\boldsymbol{P_\mu} = [\boldsymbol{e_1}]^T$. Thus:

$$S_{\mu,i}(\omega) = \{(\hat{\boldsymbol{\mu}}_{\boldsymbol{i}} - \boldsymbol{\mu_i})^T((\boldsymbol{A}^T\boldsymbol{A})^{-1})^{-1}_{11}(\hat{\boldsymbol{\mu}}_{\boldsymbol{i}} - \boldsymbol{\mu_i}) \leq (m+1)c_{m+1}(\omega)\sigma^2_{\epsilon_i}\}$$

And for the complete opportunity set we gain

$$S_\mu(\omega) = \{\boldsymbol{\mu} \in [\boldsymbol{\mu_0} - \gamma_i, \boldsymbol{\mu_0} + \gamma_i], i = 1, \dots, n\}$$

where

$$\mu_0 = \hat{\mu}, \quad \gamma_i = \sqrt{(m+1)(\boldsymbol{A^T A})^{-1}_{11} c_{m+1}(w)\sigma^2_{\epsilon_i}}, \;\; i = 1, \dots, n$$

The projection matrix for $\boldsymbol{V}$ is $\boldsymbol{P_v} = [\boldsymbol{e_2}, \boldsymbol{e_3}, \dots, \boldsymbol{e_{m+1}}]^T \in \boldsymbol{R}^{m\times(m+1)}$ and the uncertainty set for $\boldsymbol{V}$ is therefore:

$$S_v(\omega) = \{\boldsymbol{V} : (\hat{\boldsymbol{V}}_{\boldsymbol{0}} - \boldsymbol{V_0})^T(\boldsymbol{P_v}^T(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{P_v})^{-1}(\hat{\boldsymbol{V}}_{\boldsymbol{0}} - \boldsymbol{V_0}) \leq (m+1)c_{m+1}(\omega)\sigma^2_{\epsilon_i}\}$$

This is equivalent to the following simplified notation:

$$S_v(\omega) = \{\boldsymbol{V} : \boldsymbol{V_0} + \boldsymbol{W}, ||\boldsymbol{W_i}||_G \leq \rho_i, i = 1, \dots, n\}$$

where

$$\begin{aligned} \boldsymbol{V}_0 &= [\hat{\boldsymbol{V}}_{\boldsymbol{1}} \dots \hat{\boldsymbol{V}}_{\boldsymbol{n}}] \\ \rho_i &= \sqrt{(m+1)c_{m+1}(\omega)\sigma^2_{\epsilon_i}}, \;\; i = 1, \dots, n \\ ||\boldsymbol{u}||_G &= \sqrt{\boldsymbol{u^T G u}} \text{ with } \boldsymbol{G} = (\boldsymbol{P_v}^T(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{P_v})^{-1} \end{aligned}$$

For actual calculation of the matrix $\boldsymbol{G}$ it is not necessary to inverse $(\boldsymbol{A}^T\boldsymbol{A})$. To see this, let us consider the following:

$$\boldsymbol{A^T A} = \begin{bmatrix} \boldsymbol{1^T} \\ \boldsymbol{B} \end{bmatrix} \begin{bmatrix} \boldsymbol{1} & \boldsymbol{B^T} \end{bmatrix} = \begin{bmatrix} \boldsymbol{1^T 1} & (\boldsymbol{B1})^T \\ \boldsymbol{B1} & \boldsymbol{BB^T} \end{bmatrix}$$

Now we calculate the inverse matrix of $(\boldsymbol{A^T A})$ but concentrate only on the relevant right lower submatrix:

$$\begin{array}{c|c}
\begin{bmatrix} k & (\boldsymbol{B1})^T \\ \boldsymbol{B1} & \boldsymbol{BB}^T \end{bmatrix} & \begin{bmatrix} 1 & \boldsymbol{0}^T \\ \boldsymbol{0} & \boldsymbol{I}_{m\times m} \end{bmatrix} \\
\begin{bmatrix} 1 & \frac{1}{k}(\boldsymbol{B1})^T \\ \boldsymbol{B1} & \boldsymbol{BB}^T \end{bmatrix} & \begin{bmatrix} \frac{1}{k} & \boldsymbol{0}^T \\ \boldsymbol{0} & \boldsymbol{I}_{m\times m} \end{bmatrix} \\
\begin{bmatrix} \boldsymbol{1} & \frac{1}{k}(\boldsymbol{B1})^T \\ \boldsymbol{0} & \boldsymbol{BB}^T - \frac{1}{k}(\boldsymbol{B1})(\boldsymbol{B1})^T \end{bmatrix} & \begin{bmatrix} \frac{1}{k} & \boldsymbol{0}^T \\ -\frac{1}{k}(\boldsymbol{B1}) & \boldsymbol{I}_{m\times m} \end{bmatrix} \\
\begin{bmatrix} 1 & \boldsymbol{0}^T \\ \boldsymbol{0} & \boldsymbol{BB}^T - \frac{1}{k}(\boldsymbol{B1})(\boldsymbol{B1})^T \end{bmatrix} & \begin{bmatrix} \frac{1}{k} & -\boldsymbol{XI} \\ -\frac{1}{k}(\boldsymbol{B1}) & \boldsymbol{I}_{m\times m} \end{bmatrix} \\
\begin{bmatrix} \frac{1}{k} & \boldsymbol{0}^T \\ \boldsymbol{0} & \boldsymbol{I}_{m\times m} \end{bmatrix} & \begin{bmatrix} \frac{1}{l} & -\boldsymbol{XI} \\ -\frac{1}{k}(\boldsymbol{B1}) & (\boldsymbol{BB}^T - \frac{1}{k}(\boldsymbol{B1})(\boldsymbol{B1})^T)^{-1} \end{bmatrix}
\end{array}$$

Thus:

$$\boldsymbol{G} = (\boldsymbol{P_v}(\boldsymbol{A^T}\boldsymbol{A})^{-1}\boldsymbol{P_v^T})^{-1} = \boldsymbol{BB^T} - \frac{1}{p}(\boldsymbol{B1})(\boldsymbol{B1})^T$$

With this we found two uncertainty sets corresponding to $\boldsymbol{\mu}$ and $\boldsymbol{V}$ starting with an ω-confidence region for the combined vector $\boldsymbol{x}$. However, solving investment problems we would like to start with two $\hat{\omega}$-sets of the two vectors $\boldsymbol{\mu}$ and $\boldsymbol{V}$ and combine them to one joint confidence region of $\boldsymbol{x}$. This can be done analog to the argumentation above. Then we have two $\mathcal{F}$-distributed variables of the form:

$$\mathcal{Y} = \frac{1}{Js_i^2}(\boldsymbol{P\hat{x}_i} - \boldsymbol{Px_i})^T(\boldsymbol{P}(\boldsymbol{A^T A})^{-1}\boldsymbol{P}^T)^{-1}(\boldsymbol{P\hat{x}_i} - \boldsymbol{Px_i})$$

with J degrees of freedom in the numerator and $k-m-1$ degrees of freedom in the denominator with:

$$P(\mathcal{Y} \leq c_J(\hat{\omega})) = \hat{\omega}$$

We know already that the projection matrices are $\boldsymbol{P}_\mu = [\boldsymbol{e_1}]^T$ and $\boldsymbol{P}_V = [\boldsymbol{e_2, \ldots, e_{m+1}}]^T$. Thus the $\hat{\omega}$-confidence regions of $\boldsymbol{\mu}$ and $\boldsymbol{V}$ are:

$$S_\mu(\omega) = \{\boldsymbol{\mu} \in [\boldsymbol{\mu_0} - \gamma_i, \boldsymbol{\mu_0} + \gamma_i], i = 1, \ldots, n\} \tag{3.2}$$

where

$$\mu_0 = \bar{\mu}, \quad \gamma_i = \sqrt{(m+1)(\boldsymbol{A^T A})_{11}^{-1} c_1(w)\sigma_{\epsilon_i}^2}, \;\; i = 1, \ldots, n \tag{3.3}$$

and:

$$S_v(\omega) = \boldsymbol{V} : \boldsymbol{V_0} + \boldsymbol{W}, ||\boldsymbol{W_i}||_G \leq \rho_i, i = 1, \ldots, n \tag{3.4}$$

where

$$\begin{aligned}
\boldsymbol{V}_0 &= [\boldsymbol{\bar{V}_1 \ldots \bar{V}_n}] && (3.5)\\
\rho_i &= \sqrt{(m+1)c_{m+1}(\omega)\sigma_{\epsilon_i}^2}, \;\; i = 1, \ldots, n && (3.6)\\
\boldsymbol{G} &= \boldsymbol{G} = (\boldsymbol{P_v}(\boldsymbol{A^T A})^{-1}\boldsymbol{P_v^T})^{-1} = \boldsymbol{BB^T} - \frac{1}{p}(\boldsymbol{B1})(\boldsymbol{B1})^T && (3.7)\\
||\boldsymbol{u}||_G &= \sqrt{\boldsymbol{u^T G u}} && (3.8)
\end{aligned}$$

The joint confidence region is then

$$S_x = \{\boldsymbol{x} :\in S_\mu \times S_v\}$$

With a probability not less than:

$$\begin{aligned}
P(\boldsymbol{x} \in S_\mu \times S_v) &= 1 - P([\boldsymbol{\mu}, \boldsymbol{V}] \notin S_\mu \times S_v)\\
&\geq 1 - P(\boldsymbol{\mu} \notin S_\mu(\hat{\omega})) - P(\boldsymbol{V} \notin S_v(\hat{\omega}))\\
&= 1 - 2(1-\hat{\omega}) = 2\hat{\omega} - 1
\end{aligned}$$

Chapter 4

Robust Investment Problems

Using uncertainty sets instead of fixed parameters forces us to reformulate the investment problems introduced in Chapter 2. Goldfarb and Iyengar [2] tackle this problems by reformulating the tasks as robust problems. The underlying market model for their robust formulation is a multi-index model with m market factors driving the returns $\boldsymbol{r} \in \mathrm{R}^{\boldsymbol{n}}$ of the n assets in the market:

$$\boldsymbol{r} = \boldsymbol{\mu} - \boldsymbol{V^T f} + \boldsymbol{\epsilon}$$

where	$\boldsymbol{\mu} \in \mathrm{R}^n$	is the vector of mean returns of the assets,
	$\boldsymbol{f} \in \mathrm{R}^m$	is the vector of the mean index returns, with $\boldsymbol{f} \sim N(0, F)$
	$\boldsymbol{V} \in \mathrm{R}^{m \times n}$	is the beta-matrix for all assets and indices,
	$\boldsymbol{\epsilon} \in R^n$	is the vector of residuals, with $\boldsymbol{\epsilon} \sim N(0, D)$

$\boldsymbol{x} \sim N(\boldsymbol{mean}, \boldsymbol{cov})$ denotes that $\boldsymbol{x}$ is a multivariate Normal random variables with mean vector $\boldsymbol{mean}$ and covariance matrix $\boldsymbol{cov}$. Additional Goldfarb and Iyengar [2] use the assumption of a standard multi index model, i.e.

$E[\epsilon_i f_j] = 0$	the residuals are independent of the factor returns
$\boldsymbol{F} \succ 0$	the covariance matrix of the indices is strictly positive definite
$E[\epsilon_i \epsilon_j] = 0$	the residuals are independent from each other, i.e. $\boldsymbol{D} = \mathbf{diag}(d) \succeq 0$

Therefore the probability distribution of the return vector $\boldsymbol{r} \in \boldsymbol{R^n}$ is given by:

$$\boldsymbol{r} \sim N(\boldsymbol{\mu}, \boldsymbol{V^T F V} + \boldsymbol{D})$$

Instead of ignoring the fact that the input parameters $\boldsymbol{\mu}, \boldsymbol{V}, \boldsymbol{D}$ can only be estimated and never known exactly Goldberg and Iyengar [2] work with the uncertainty sets derived in Section 3.4:

$$S_\mu = \{\boldsymbol{\mu} \mid \boldsymbol{\mu_0} + \boldsymbol{\zeta},\ |\zeta_i| \leq \gamma_i, i = 1, \dots, n\} \tag{4.1}$$

$$S_v = \{\boldsymbol{V} \mid \boldsymbol{V} = \boldsymbol{V_0} + \boldsymbol{W}, \|\boldsymbol{W_i}\|_g \leq \rho_i,\ i = 1, \dots, n\} \tag{4.2}$$

$$S_d = \{\boldsymbol{D} \mid \boldsymbol{D} = \mathbf{diag}(\boldsymbol{d}),\ d_i \in [\underline{d_i}, \overline{d_i}], i = 1, \dots, n\} \tag{4.3}$$

where $\underline{d_i} \leq \overline{d_i}, \rho_i, \gamma_i \in R_+$ and $\boldsymbol{W_i}$ is the i-th column vector of the matrix $\boldsymbol{W}$ and $||\boldsymbol{u}||_g = \sqrt{\boldsymbol{u^T G u}}$ is a norm defined by a given symmetric and positive definite matrix $\boldsymbol{G}$.

Goldberg also formulates an approach with an uncertainty set for the matrix $\boldsymbol{F}$ (see section 6 [2]) but in this master thesis I will concentrate on his first approach (see section 2 [2]) and assume that the covariance matrix $\boldsymbol{F}$ of the index returns is stable and exact.

So let's see how this new view on the input parameters changes the problem formulation for the three investment problems. Until now the **minimum variance portfolio** was the portfolio $\boldsymbol{\phi}$ of all portfolios with an expected return of at least α which had the lowest risk. Now we no longer assume the input parameters to be known exactly, but to vary within given uncertainty sets. Hence, the most conservative or risk averse approach is to minimize the worst case variance subject to the condition that the worst case expected return must not be less then α. This is called the robust formulation of the minimum variance problem. So our optimization problem for the minimum variance problem changes as follows:

markowitz formulation		**robust formulation**	
minimize	$\boldsymbol{Var}[r_\phi]$	minimize	$\max_{\{\boldsymbol{V} \in S_v, \boldsymbol{D} \in S_D\}} \boldsymbol{Var}[r_\phi]$
subject to	$\boldsymbol{E}[r_\phi] \geq \alpha$	subject to	$\min_{\{\boldsymbol{\mu} \in S_\mu\}} \boldsymbol{E}[r_\phi] \geq \alpha$
	$\boldsymbol{1}^T \boldsymbol{\phi} = 1$		$\boldsymbol{1}^T \boldsymbol{\phi} = 1$

Analogous is the aim of the robust formulation of the **maximum return problem** to maximize the worst case expected return if the worst case risk allowed is λ:

markowitz formulation		**robust formulation**	
maximize	$\boldsymbol{E}[r_\phi]$	maximize	$\min_{\{\boldsymbol{\mu} \in S_\mu\}} \boldsymbol{E}[r_\phi]$
subject to	$\boldsymbol{Var}[r_\phi] \leq \lambda$	subject to	$\max_{\{\boldsymbol{V} \in S_v, \boldsymbol{D} \in S_D\}} \boldsymbol{Var}[r_\phi] \leq \lambda$
	$\boldsymbol{1}^T \boldsymbol{\phi} = 1$		$\boldsymbol{1}^T \boldsymbol{\phi} = 1$

And last but not least is the **robust maximum Sharpe ratio problem** the task to maximize the worst case ratio of the return above the risk-free rate to the variance of the return:

markowitz formulation		**robust formulation**	
maximize	$\frac{\boldsymbol{E}[r_\phi] - r_f}{\sqrt{\boldsymbol{Var}[r_\phi]}}$	maximize	$\min_{\{\boldsymbol{\mu} \in S_\mu, \boldsymbol{D} \in S_d\}} \frac{\boldsymbol{E}[r_\phi] - r_f}{\sqrt{\boldsymbol{Var}[r_\phi]}}$
subject to	$\boldsymbol{1}^T \boldsymbol{\phi} = 1$	subject to	$\boldsymbol{1}^T \boldsymbol{\phi} = 1$

To find a numerical solution of these problems Goldfarb and Iyengar [2] suggest to represent them as second order cone problems. The next section will give the definition of this problem class.

4.1 Definition of Second Order Cone Problems

According to Lobo [3] second-order cone programs (SOCP) are optimization problems of the following form

$$\begin{array}{ll} \text{minimize} & \boldsymbol{f}^T \boldsymbol{x} \\ \text{subject to} & ||\boldsymbol{A_i x} + \boldsymbol{b_i}|| \leq \boldsymbol{c_i}^T \boldsymbol{x} + d_i \quad i = 1, \ldots, N \end{array}$$

where $x \in R^n$ is the variable that is to be optimized. The other parameters are known constants of the following dimensions:

$$\boldsymbol{f} \in \mathrm{R}^n, \; \boldsymbol{A_i} \in \mathrm{R}^{(n_i-1)\times n}, \; \boldsymbol{b_i} \in \mathrm{R}^{n_i-1}, \; \boldsymbol{c_i} \in R^n \text{ and } d_i \in R$$

$||.||$ is the standard Euclidean norm, i.e., $||u|| = (u^T u)^{1/2}$. The constraints

$$||A_i x + b_i|| \leq ci^T x + d_i$$

are called second-order cone constraints of dimension n_i, because the set of points described by them are the inverse images of unit second order cones under an affine mapping. To see this, recall the definition of a standard second order cone of dimension $k \geq 1$ from vector analysis:

$$Cone_k = \left\{ \left[\begin{array}{c} A \\ c \end{array} \right] \middle| \; A \in R^{k-1}, \; c \in R, ||A|| \leq c \right\}$$

The special case $k = 1$ is defined as

$$Cone_1 = \{c \mid c \in R, 0 \leq c\}$$

The relationship between a second-order cone constraint and the corresponding cone of the same dimension is given by:

$$||\boldsymbol{A_i x} + \boldsymbol{b_i}|| \leq \boldsymbol{c_i}^T x + d_i \Longleftrightarrow \left[\begin{array}{c} \boldsymbol{A_i} \\ \boldsymbol{c_i}^T \end{array} \right] x + \left[\begin{array}{c} \boldsymbol{b_i} \\ d_i \end{array} \right] \in C_{n_i} b$$

Cones are known to be convex. There exists an affine mapping between the cones to the sets of the constraints. Therefore the constraints define a convex set. As the objective function of the SOCPs is also convex, they are convex programming problems.

4.2 The minimum variance problem represented as SOCP

We will now show at the example of the minimum variance problem that the robust investment problems are indeed equivalent to second order cone problems.

4.3 Rewriting the constraints on S_d and S_μ

As we have seen in Section 4 the robust formulation of the minimum variance problem is given by:

$$\begin{array}{ll} \text{minimize} & \max_{\{\boldsymbol{V} \in S_v\}} \{\boldsymbol{\phi^T V^T F V \phi}\} + \max_{\{\boldsymbol{D} \in S_d\}} \{\boldsymbol{\phi^T D \phi}\} \\ \text{subject to} & \min_{\{\boldsymbol{\mu} \in S_\mu\}} \boldsymbol{\mu^T \phi} \geq \alpha, \\ & \boldsymbol{1^T \phi} = 1 \end{array} \tag{4.4}$$

To reformulate the minimum variance problem as a SOCP we have to get rid of the minimum and maximum terms in the objective function and the constraints. We can achieve this for the objective function by introducing the auxiliary variables $\nu, \delta \in \mathrm{R}$ as follows:

$$\begin{array}{ll} \text{minimize} & \nu + \delta \\ \text{subject to} & \max_{\{\boldsymbol{V} \in S_v\}} \boldsymbol{\phi}^{\boldsymbol{T}} \boldsymbol{V}^{\boldsymbol{T}} \boldsymbol{F} \boldsymbol{V} \boldsymbol{\phi} \leq \nu \\ & \max_{\{\boldsymbol{D} \in S_d\}} \{\boldsymbol{\phi}^{\boldsymbol{T}} \boldsymbol{D} \boldsymbol{\phi}\} \leq \delta \\ & \min_{\{\boldsymbol{\mu} \in S_\mu\}} \boldsymbol{\mu}^{\boldsymbol{T}} \boldsymbol{\phi} \geq \alpha, \\ & \boldsymbol{1}^{\boldsymbol{T}} \boldsymbol{\phi} = 1 \end{array}$$

Let us reconsider the definitions of the uncertainty sets S_d and S_μ:

$$\begin{array}{rcl} S_d & = & \{\boldsymbol{D} \mid \boldsymbol{D} = \mathbf{diag}(\boldsymbol{d}),\ d_i \in [\underline{d_i}, \overline{d_i}], i = 1, \ldots, n\} \\ S_\mu & = & \{\boldsymbol{\mu} \mid \boldsymbol{\mu_0} + \boldsymbol{\zeta},\ |\zeta_i| \leq \gamma_i, i = 1, \ldots, n\} \end{array}$$

Thus the above constraints on S_d and S_μ can be rewritten as:

$$\begin{array}{lclcl} \max_{\{\boldsymbol{D} \in S_d\}} \{\boldsymbol{\phi}^{\boldsymbol{T}} \boldsymbol{D} \boldsymbol{\phi}\} & = & \sum_{i=1}^n \phi_i^2 \max d_i & = & \sum_{i=1}^n \phi_i^2 \bar{d}_i = \boldsymbol{\phi}^{\boldsymbol{T}} \bar{\boldsymbol{D}} \boldsymbol{\phi} \\ \min_{\{\boldsymbol{\mu} \in S_\mu\}} \boldsymbol{\mu}^{\boldsymbol{T}} \boldsymbol{\phi} & = & \boldsymbol{\mu_0^T} \boldsymbol{\phi} - \boldsymbol{\gamma}^{\boldsymbol{T}} |\boldsymbol{\phi}| & & \end{array}$$

Since the covariance matrix $\boldsymbol{F}$ is strictly positive definite, $||\boldsymbol{u}||_f = \sqrt{\boldsymbol{u}^{\boldsymbol{T}} \boldsymbol{F} \boldsymbol{u}}$ defines a norm on R^m. Thus the minimization problem (4.4) can be rewritten as:

$$\begin{array}{ll} \text{minimize} & \nu + \delta \\ \text{subject to} & \max_{\{\boldsymbol{V} \in S_v\}} ||\boldsymbol{V} \boldsymbol{\phi}||_f^2 \leq \nu \\ & \boldsymbol{\phi}^{\boldsymbol{T}} \bar{\boldsymbol{D}} \boldsymbol{\phi} \leq \delta \\ & \boldsymbol{\mu_0^T} \boldsymbol{\phi} - \boldsymbol{\gamma}^{\boldsymbol{T}} |\boldsymbol{\phi}| \geq \alpha, \\ & \boldsymbol{1}^{\boldsymbol{T}} \boldsymbol{\phi} = 1 \end{array}$$

Substitution of the maximization term on the uncertainty set S_v is a bit more work.

4.4 Rewriting the constraints on S_v

Reconsider the definition of S_v:

$$S_v = \{\boldsymbol{V} \mid \boldsymbol{V} = \boldsymbol{V_0} + \boldsymbol{W},\ ||\boldsymbol{W_i}||_g \leq \rho_i,\ i = 1, \ldots, n\}$$

where $||\boldsymbol{u}||_g = \sqrt{\boldsymbol{u}^{\boldsymbol{T}} \boldsymbol{G} \boldsymbol{u}}$ is the norm on R^m defined by the strictly positive definite matrix G. Thus to find the worst case variance for a fixed portfolio $\boldsymbol{\phi}$ we have to solve the following optimization problem:

$$\begin{array}{ll} \text{maximize} & ||(\boldsymbol{V_0} + \boldsymbol{W}) \boldsymbol{\phi}||_f^2 \\ \text{subject to} & ||\boldsymbol{W_i}||_g \leq \rho_i,\ i = 1, \ldots, n \end{array} \tag{4.5}$$

Lemma 5 *The optimization problem (4.5) is equivalent to:*

$$\begin{array}{ll} \textit{maximize} & ||\boldsymbol{V_0\phi} + \boldsymbol{W\phi}||_f^2 \\ \textit{subject to} & ||\boldsymbol{W\phi}||_g \leq \boldsymbol{\rho^T}|\boldsymbol{\phi}| \end{array} \tag{4.6}$$

Proof:

For all $\boldsymbol{W} \in \{\boldsymbol{W} \in R^{m\times n}|\ \ ||\boldsymbol{W_i}||_g \leq \rho_i, i = 1, \dots, n\}$ the following inequality holds:

$$||\boldsymbol{W\phi}||_g = \left\|\sum_{i=1}^{n} \phi_i \boldsymbol{W_i}\right\|_g \leq \sum_{i=1}^{n} |\phi_i|||\boldsymbol{W_i}||_g \leq \sum_{i=1}^{n} \rho_i|\phi_i| = \boldsymbol{\rho^T}|\boldsymbol{\phi}|$$

So that $\{\boldsymbol{W} \in R^{m\times n}|\ \ ||\boldsymbol{W_i}||_g \leq \rho_i, i = 1, \dots, n\} \subset \{\boldsymbol{W} \in R^{m\times n}|\ \ ||\boldsymbol{W\phi}||_g \leq \boldsymbol{\rho^T}|\boldsymbol{\phi}|\}$, i.e. the optimal solution value of (4.5) is also a possible solution of (4.6) and the maximum of (4.6) equal or greater than that of (4.5).

Let $\boldsymbol{W^*}$ denote the optimal solution of (4.6). Then the two optimization problems are equivalent if and only if there exists a $\tilde{\boldsymbol{W}}$ such that $\tilde{\boldsymbol{W}}\boldsymbol{\phi} = \boldsymbol{W^*\phi}$ and $\bar{\boldsymbol{W}}$ is in the feasible set of (4.6).

Because the objective function $||V_0\phi + W\phi||$ is convex we know from analysis, that the optimal solution W^* of (4.6) is found on the boundary of the feasible set. Therefore we know that: $||\boldsymbol{W^*\phi}||_{\boldsymbol{g}} = \boldsymbol{\rho^T}|\boldsymbol{\phi}|$. Now define :

$$\tilde{\boldsymbol{W}}_{\boldsymbol{i}} = \begin{cases} \frac{\phi_i}{|\phi_i|}\frac{\phi_i}{\boldsymbol{\rho^T}|\boldsymbol{\phi}|}\boldsymbol{W^*\phi} & \text{if } \phi_i \neq 0 \\ \frac{\phi_i}{\boldsymbol{\rho^T}|\boldsymbol{\phi}|}\boldsymbol{W^*\phi} & \text{otherwise} \end{cases}$$

Then

$$\tilde{\boldsymbol{W}}\phi = \sum_{i=1}^{n}\phi_i W_i = \sum_{i=1}^{n}\phi_i \left\{ \begin{array}{ll} \frac{\phi_i}{|\phi_i|}\frac{\phi_i}{\boldsymbol{\rho^T}|\boldsymbol{\phi}|}\boldsymbol{W^*\phi} & \text{if } \phi_i \neq 0 \\ \frac{\phi_i}{\boldsymbol{\rho^T}|\boldsymbol{\phi}|}\boldsymbol{W^*\phi} & \text{otherwise} \end{array} \right\} = \frac{\boldsymbol{W^*\phi}}{\boldsymbol{\rho^T}|\boldsymbol{\phi}|}\sum_{i=1}^{n}\rho_i|\phi_i| = \boldsymbol{W^*\phi} \tag{4.7}$$

and

$$||\tilde{\boldsymbol{W}}_{\boldsymbol{i}}||_g = \left\{ \begin{array}{ll} \frac{|\phi_i|}{|\phi_i|}\frac{\rho_i}{\boldsymbol{\rho^T}|\boldsymbol{\phi}|}||\boldsymbol{W^*\phi}||_g & \text{if } \phi_i \neq 0 \\ \frac{\rho_i}{\rho^T|\phi|}||\boldsymbol{W^*\phi}||_g & \text{otherwise} \end{array} \right\} = \frac{\rho_i|\phi|}{\boldsymbol{\rho^T}|\boldsymbol{\phi}|}\boldsymbol{\rho^T}|\boldsymbol{\phi}| = \rho_i|\phi| \qquad \square$$

Thus, the constraint for the worst case variance of a fixed portfolio ϕ is of the form:

$$\max_{\{y:||\boldsymbol{y}||_g\leq r\}} ||\boldsymbol{y_0} + \boldsymbol{y}||_f^2 \leq \nu \tag{4.8}$$

where $\boldsymbol{y_0} = \boldsymbol{V_0\phi}, \boldsymbol{y} = \boldsymbol{W\phi}$ and $r = \rho^T|\phi|$.

In the next step toward a second order cone program we will proof a lemma that allows us to replace a maximization problem of the form (4.8) by a set of linear equalities, inequalities and so called restricted hyperbolic constraints of the form: $\boldsymbol{a^T a} \leq bc, \boldsymbol{a} \in \mathrm{R}^n, b, c \in \mathrm{R}$ and $a, b \geq 0$.

Lemma 6 *Let $r, \nu > 0, \boldsymbol{y_0}, \boldsymbol{y} \in R^m$ and $\boldsymbol{F}, \boldsymbol{G} \in R^{m\times m}$ be positive definite matrices. Then the constraint*

$$\max_{\{y:\|\boldsymbol{y}\|_g \leq r\}} \|\boldsymbol{y_0} + \boldsymbol{y}\|_f^2 \leq \nu \tag{4.9}$$

holds if and only if there exist $\tau, \sigma > 0 \in R$, and $t \in R_+^m$ that satisfy

$$\begin{aligned} \nu &\geq \tau + 1^T t \\ \sigma &\leq \frac{1}{\lambda_{\max}(H)} \\ r^2 &\leq \sigma\tau \\ w_i^2 &\leq (1 - \sigma\lambda_i)t_i, \;\; i = 1, \ldots, m \end{aligned}$$

where $\boldsymbol{H} := \boldsymbol{G^{-1/2} F G^{-1/2}}$ with the spectral decomposition $\boldsymbol{H} = \boldsymbol{Q\Delta Q^T}, \boldsymbol{\Delta} = diag(\lambda_i)$ and $w = Q^T H^{1/2} G^{1/2} y_0$

Proof:

$r \geq 0$ by defininition. Therefore substituting $\boldsymbol{y} = \boldsymbol{r\bar{y}}$ yields

$$\begin{aligned} \|\boldsymbol{y_0} + \boldsymbol{y}\|_f^2 &= \boldsymbol{(y_0 + r\bar{y})^T F (y_0 + r\bar{y}) = y_0^T F y_0 + 2r y_0^T F \bar{y} + r^2 \bar{y}^T F \bar{y}} \\ \|\boldsymbol{y}\|_g &= \|r\boldsymbol{\bar{y}}\|_g = r\|\boldsymbol{\bar{y}}\|_g \end{aligned}$$

Thus:

$$\|y\|_g \leq r \Longleftrightarrow \|\bar{y}\|_g \leq 1 \Longleftrightarrow \|\bar{y}\|_g^2 \leq 1 \Longleftrightarrow 1 - \boldsymbol{\bar{y}^T G \bar{y}} \geq 0$$

Therefore (4.9) is equivalent to

$$(\nu - y_0^T F y_0) - 2r y_0^T F \bar{y} - r^2 \bar{y}^T F \bar{y} \geq 0 \quad \text{for all} \quad \bar{y} \;\text{ such that } 1 - \bar{y}^T G \bar{y} \geq 0$$

Now we've got two quadratic functions on which we can can use the $S-$ procedure (see Goldfarb and Iyengar [2]):

Lemma 7 *(S-procedure) Let $P_i(\boldsymbol{x}), i = 1, \ldots, k$ be quadratic functions of $\boldsymbol{x} \in \boldsymbol{R^n}$:*

$$P_i(x) = \boldsymbol{x^T A_i x + 2b_i^T x} + c_i, \; i = 0, \ldots, k$$

Then $P_0(\boldsymbol{x}) \geq 0$ for all x such that $P_i(\boldsymbol{x}) \geq 0, i = 1, \ldots, k$, if there exist $\tau_i \geq 0$ such that :

$$\begin{bmatrix} c_0 & b_0^T \\ b_0 & A_0 \end{bmatrix} - \sum_{i=1}^{p} \tau_i \begin{bmatrix} c_i & b_i^T \\ b_i & A_i \end{bmatrix} \succeq 0$$

Moreover, if $k = 1$ then the converse holds if there exists $\boldsymbol{x_0}$ such that $P_1(\boldsymbol{x_0}) > 0$.

Here $k = 1$ and the two quadratic functions are given by:

$$\begin{aligned} P_0(\boldsymbol{\bar{y}}) &= \boldsymbol{\bar{y}^T}(-r^2 \boldsymbol{F})\boldsymbol{\bar{y}} + 2(-r\boldsymbol{y_0^T F})\boldsymbol{\bar{y}} + (\nu - \boldsymbol{y_0^T F y_0}) \\ P_1(\boldsymbol{\bar{y}}) &= \boldsymbol{\bar{y}^T(-G)\bar{y}} + 1 \end{aligned}$$

Since $P_1(\mathbf{0}) = 1 > 0$ the S-procedure tells us that $P_0(\bar{\boldsymbol{y}}) \geq 0$ for all $\bar{\boldsymbol{y}}$ such that $P_1(\bar{\boldsymbol{y}}) \geq 0$ if and only if there exists a $\tau \geq 0$ such that

$$\boldsymbol{M} = \begin{bmatrix} \nu - \boldsymbol{y_0^T F y_0} & -r\boldsymbol{y_0^T F} \\ -r\boldsymbol{F y_0} & -r^2\boldsymbol{F} \end{bmatrix} - \tau \begin{bmatrix} 1 & \boldsymbol{0^T} \\ \boldsymbol{0} & -\boldsymbol{G} \end{bmatrix} \succeq 0$$

By definition of $\boldsymbol{H}, \boldsymbol{Q}, \boldsymbol{\Delta}$ and $\boldsymbol{w}$:

$$\boldsymbol{y_0^T F y_0} \quad = \quad \boldsymbol{y_0^T G^{\frac{1}{2}} H G^{\frac{1}{2}} y_0} = \boldsymbol{y_0^T G^{\frac{1}{2}} H^{\frac{1}{2}} Q Q^T H^{\frac{1}{2}} G^{\frac{1}{2}} y_0} = \boldsymbol{w^T w}$$

So $\boldsymbol{M}$ can alternatively be written as:

$$\boldsymbol{M} = \begin{bmatrix} \nu - \tau - \boldsymbol{w^T w} & -r\boldsymbol{y_0^T F} \\ -r\boldsymbol{F y_0} & \tau\boldsymbol{G} - r^2\boldsymbol{F} \end{bmatrix} \succeq 0$$

Define $\bar{\boldsymbol{M}} = \boldsymbol{Z M Z^T}$, where

$$\boldsymbol{Z} = \begin{bmatrix} 1 & \boldsymbol{0^T} \\ \boldsymbol{0} & \boldsymbol{Q^T G^{-\frac{1}{2}}} \end{bmatrix}$$

then the Matrix $\boldsymbol{M} \succeq 0 \Longleftrightarrow \bar{\boldsymbol{M}} \succeq 0$. So let's calculate the entries of $\bar{\boldsymbol{M}}$:

$$\begin{aligned} \boldsymbol{M} &= \begin{bmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0^T} & \boldsymbol{Q^T G^{-\frac{1}{2}}} \end{bmatrix} \begin{bmatrix} \nu - \tau - \boldsymbol{w^T w} & -r\boldsymbol{y_0^T F} \\ -r\boldsymbol{F y_0} & \tau\boldsymbol{G} - r^2\boldsymbol{F} \end{bmatrix} \begin{bmatrix} 1 & \boldsymbol{0^T} \\ \boldsymbol{0} & \boldsymbol{G^{-\frac{1}{2}} Q} \end{bmatrix} \\ &= \begin{bmatrix} \nu - \tau - \boldsymbol{w^T w} & -r\boldsymbol{y_0^T F} \\ -r\boldsymbol{Q^T G^{-\frac{1}{2}} F y_0} & \tau\boldsymbol{Q^T G^{\frac{1}{2}}} - r^2\boldsymbol{Q^T G^{-\frac{1}{2}} F} \end{bmatrix} \begin{bmatrix} 1 & \boldsymbol{0^T} \\ \boldsymbol{0} & \boldsymbol{G^{-\frac{1}{2}} Q} \end{bmatrix} \\ &= \begin{bmatrix} \nu - \tau - \boldsymbol{w^T w} & -r\boldsymbol{y_0^T F G^{-\frac{1}{2}} Q} \\ -r\boldsymbol{Q^T G^{-\frac{1}{2}} F y_0} & \tau\boldsymbol{Q^T Q} - r^2\boldsymbol{Q^T G^{-\frac{1}{2}} F G^{-\frac{1}{2}} Q} \end{bmatrix} \end{aligned}$$

Using the definitions of $\boldsymbol{H}, \boldsymbol{Q}, \boldsymbol{\Delta}$ and $\boldsymbol{QQ^T} = \boldsymbol{Q^T Q} = \boldsymbol{I}$ we have:

$$\begin{aligned} &\boldsymbol{Q^T G^{-\frac{1}{2}} F y_0} = \boldsymbol{Q^T G^{-\frac{1}{2}} F G^{-\frac{1}{2}} G^{\frac{1}{2}} y_0} = \boldsymbol{Q^T H G^{\frac{1}{2}} y_0} = \boldsymbol{Q^T H^{\frac{1}{2}} Q Q^T H^{\frac{1}{2}} G^{\frac{1}{2}} y_0} = \boldsymbol{\Delta^{\frac{1}{2}} w} \\ &\boldsymbol{Q^T G^{-\frac{1}{2}} F G^{-\frac{1}{2}} Q} = \boldsymbol{Q^T H Q} = \boldsymbol{Q^{-1} H (Q^T)^{-1}} = \boldsymbol{\Delta} \end{aligned}$$

So we end up with:

$$\bar{\boldsymbol{M}} = \begin{bmatrix} \nu - \tau - \boldsymbol{w^T w} & -r\boldsymbol{w^T \Delta^{\frac{1}{2}}} \\ -r\boldsymbol{\Delta^{\frac{1}{2}} w} & \tau\boldsymbol{I} - r^2\boldsymbol{\Delta} \end{bmatrix}$$

From linear Algebra (see Lorenz [4] p.70) we know that a matrix A is positive definite if and only if

$$|A_{S,S}| \geq 0 \quad \text{for all submatrices } A_{S,S} \text{ with } S \; in \begin{pmatrix} n \\ k \end{pmatrix}$$

So let's have a closer look inside $\bar{\boldsymbol{M}}$:

$$\bar{\boldsymbol{M}} = \begin{bmatrix} \nu - \tau - w^T w & -rw_1\sqrt{\lambda_1} & \dots & \dots & -rw_n\sqrt{\lambda_n} \\ -rw_1\sqrt{\lambda_1} & \tau - r^2\lambda_1 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & 0 & 0 \\ \vdots & 0 & 0 & \ddots & 0 \\ -rw_n\sqrt{\lambda_n} & 0 & \dots & 0 & \tau - r^2\lambda_n \end{bmatrix}$$

All submatrices not containing the first row and column are of the form $\mathbf{diag}(\tau - r^2\lambda_i)$ with $i \in S$. Therefore for $\bar{\boldsymbol{M}}$ to be positive definite the following must hold:

$$\tau - r^2\lambda_i \geq 0 \; \forall \, i = 1, \ldots, n \iff \tau \geq r^2\lambda_{\max}(H)$$

All submatrices including the first row and column have the same shape as matrice $\bar{M}$ itself. So extracting $|A_{S,S}|$ for all $S : 1 \in S$ by the first row yields:

$$\begin{aligned}
|A_{S,S}| &= (\nu - \tau - \boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{w})|\tau\boldsymbol{I} - r^2\boldsymbol{\Delta}| \\
&+ \textstyle\sum_{i\in S}(-1)^{i+2}(-rw_i\sqrt{\lambda_i}) \begin{vmatrix} -rw_1\sqrt{\lambda_1} & \tau - r^2\lambda_1 & 0 & \ldots & 0 \\ \vdots & 0 & \ddots & 0 & 0 \\ -rw_i\sqrt{\lambda_i} & 0 & 0 & 0 & 0 \\ \vdots & 0 & 0 & \ddots & 0 \\ -rw_n\sqrt{\lambda_n} & 0 & \ldots & 0 & \tau - r^2\lambda_n \end{vmatrix} \\
&= (\nu - \tau - \boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{w}) \textstyle\prod_{i\in S}(\tau - r^2\lambda_i) \\
&+ \textstyle\sum_{i\in S}(-1)^{i+2}(-rw_i\sqrt{\lambda_i})(-1)^{1+i}(-rw_i\sqrt{\lambda_i}) \prod_{j\in Si}(\tau - r^2\lambda_j)
\end{aligned}$$

So altogether we have for all $S : 1 \in S$

$$|A_{S,S}| = (\nu - \tau - w^Tw) \prod_{i\in S}(\tau - r^2\lambda_i) - \sum_{i\in S} r^2 w_i^2 \lambda_i \prod_{j\in S\setminus i}(\tau - r^2\lambda_j)$$

Thus for all $|A_{S,S}|$ with $1 \in S$ to be greater than zero the following constraint must hold:

$$(\nu - \tau - w^Tw)(\tau - r^2\lambda_i) - r^2\lambda_i w_i^2 \geq 0 \Longrightarrow w_i = 0 \quad \forall i \quad \text{where} \quad (\tau - r^2\lambda_i) = 0 \quad \text{as} \quad \lambda_i > 0$$

If S contains an i such that $\tau = r^2\lambda_i$ we know that $|A_{S,S}| = 0$. Thus:

$$|A_{S,S}| \geq \nu - \tau - w^Tw \prod_{\substack{i\in S \\ i:\tau\neq r^2\lambda_i}} (\tau - r^2\lambda_i) - \sum_{\substack{i:\tau\neq r^2\lambda_i \\ i\in S}} r^2 w_i^2 \lambda_i \prod_{j\in S\setminus i}(\tau - r^2\lambda_j) \geq 0$$

Dividing through $\prod_{i:\tau\neq r^2\lambda_i}(\tau - r^2\lambda_i)$ yields

$$\begin{aligned}
\nu - \tau - w^Tw - \sum_{\substack{i:\tau\neq r^2\lambda_i \\ i\in S}} \frac{r^2 w_i^2 \lambda_i}{\tau - r^2\lambda_i} &\geq \nu - \tau - \sum_{i:\tau\neq r^2\lambda_i} \frac{w_i^2(\tau - r^2\lambda_i)}{\tau - r^2\lambda_i} + \sum_{i:\tau\neq r^2\lambda_i} \frac{r^2 w_i^2 \lambda_i}{\tau - r^2\lambda_i} = \\
\nu - \tau - \sum_{i:\tau\neq r^2\lambda_i} \frac{w_i^2\tau}{\tau - r^2\lambda_i} &= \nu - \tau - \sum_{i:\tau\neq r^2\lambda_i} \frac{w_i^2}{1 - \frac{r^2}{\tau}\lambda_i} \geq 0
\end{aligned}$$

So altogether we have:

$$\max_{\{y:\|y\|_g\leq r\}} \|y_0 + y\|_f^2 \leq \nu \tag{4.10}$$

if and only if there exists a $\tau \geq 0$ such that

$$\tau \geq r^2 \lambda_{\max}(H)$$

$$w_i = 0 \quad \forall i \quad \text{where} \quad (\tau - r^2\lambda_i) = 0$$

$$\nu - \tau - \textstyle\sum_{i:\tau \neq r^2\lambda_i} \frac{w_i^2}{1-\frac{r^2}{\tau}\lambda_i} \geq 0$$

To simplify notation define $\sigma = \frac{r^2}{\tau}$:

$$r^2 = \sigma\tau \tag{4.11}$$

$$\sigma \leq \frac{1}{\lambda_{\max}(H)} \tag{4.12}$$

$$w_i = 0 \quad \forall i \quad \text{where} \quad (\tau - r^2\lambda_i) = 0 \tag{4.13}$$

$$\nu - \tau - \sum_{i:\sigma\lambda_i \neq 1} \frac{w_i^2}{1 - \sigma\lambda_i} \geq 0 \tag{4.14}$$

Let:

$$t_i = \begin{cases} \frac{w_i^2}{1-\sigma\lambda_i} & \forall i \quad \sigma\lambda_i \neq 1 \\ 0 & \forall i \quad \sigma\lambda_i = 1 \end{cases}$$

With this the constraints (4.13) and (4.14) are equivalent with:

$$w_i^2 = (1 - \sigma\lambda_i)t_i, \quad i = 1, \ldots, m \tag{4.15}$$

$$\nu \geq \tau + 1^T t \tag{4.16}$$

So to finish the proof of lemma (6) we just have to replace the equalities (4.11) and (4.15) by inequalities. This can be done as follows:

$$\exists\sigma^* \geq 0 : r^2 \leq \sigma^*\tau \Longleftrightarrow \exists\sigma \geq 0 : \sigma = \sigma^* - \frac{\sigma^*\tau - r^2}{\tau} \text{ and } r^2 = \sigma\tau$$

And for all $i = 1, \ldots, m$:

$$\exists t_i^* \in \mathrm{R}_+^m \; w_i^2 \leq (1 - \sigma\lambda_i)t_i^*, \Longleftrightarrow \exists t_i = t_i^* - \frac{t_i^*(1-\sigma\lambda_i) - w_i^2}{(1 - \sigma\lambda_i)} \text{ and } w_i^2 = (1 - \sigma\lambda_i)t_i,$$

$\square$

Using lemma (6) to reformulate the robust minimum variance problem (4.4) we get:

$$\begin{array}{ll} \text{minimize} & \nu + \delta \\ \text{subject to} & r^2 \leq \sigma\tau \\ & \sigma \leq \frac{1}{\lambda_{\max}(H)} \\ & \nu \geq \tau + 1^T t \\ & w_i^2 \leq (1 - \sigma\lambda_i)t_i, \quad I = 1, \ldots, m \\ & \boldsymbol{\phi^T \bar{D} \phi} \leq \delta \\ & \boldsymbol{\mu_0^T \phi} - \boldsymbol{\gamma^T} |\boldsymbol{\phi}| \geq \alpha, \\ & \boldsymbol{1^T \phi} = 1 \end{array}$$

4.5 Replacing the hyperbolic constraints

The only thing left to do is to rewrite the restricted hyperbolic constraints on r^2 and w_i^2 in the optimization problem (4.4) as second order cone constraints. This can be done with the help of the following lemma:

Lemma 8

$$\boldsymbol{u}^T\boldsymbol{u} \leq xy, x \geq 0, y \geq 0 \Longleftrightarrow \left\| \begin{matrix} 2\boldsymbol{u} \\ x-y \end{matrix} \right\| \leq x+y$$

Proof:

$$\begin{aligned}
\boldsymbol{u}^T\boldsymbol{u} \leq xy, x \geq 0, y \geq 0 &\implies 4\boldsymbol{u}^T\boldsymbol{u} + x^2 + y^2 - 2xy \leq x^2 + y^2 + 2xy \\
&\Longleftrightarrow \sqrt{(2\boldsymbol{u})^T(2\boldsymbol{u}) + (x-y)^2} \leq \sqrt{(x+y)^2} \\
&\Longleftrightarrow \left\| \begin{matrix} 2\boldsymbol{u} \\ x-y \end{matrix} \right\| \leq x+y \\
\sqrt{4\boldsymbol{u}^T\boldsymbol{u} + (x-y)^2} \leq x+y &\implies 4\boldsymbol{u}^T\boldsymbol{u} + x^2 + y^2 - 2xy \leq x^2 + y^2 + 2xy \wedge x+y \geq 0 \\
&\implies 4\boldsymbol{u}^T\boldsymbol{u} - 2xy \leq 2xy \wedge x+y \geq 0 \\
&\implies \boldsymbol{u}^T\boldsymbol{u} \leq xy \wedge x \geq 0 \wedge y \geq 0
\end{aligned}$$

$\square$

Therefore, the robust portfolio selection problem (4.4) is equivalent to the SOCP:

$$\begin{array}{ll}
\text{minimize} & \nu + \delta \\
\text{subject to} & \boldsymbol{w} = \boldsymbol{Q}^T\boldsymbol{H}^{\frac{1}{2}}\boldsymbol{G}^{\frac{1}{2}}\boldsymbol{V}_0\boldsymbol{\phi} \\
& \tau + \boldsymbol{1}^T\boldsymbol{t} \leq \nu \\
& \sigma\lambda_{\max}(\boldsymbol{H}) \leq 1 \\
& \left\| \begin{bmatrix} 2\boldsymbol{\rho}^T\boldsymbol{\phi} \\ \sigma - \tau \end{bmatrix} \right\| \leq \sigma + \tau \\
& \left\| \begin{bmatrix} 2w_i \\ 1 - \sigma\lambda_i - t_i \end{bmatrix} \right\| \leq 1 - \sigma\lambda_i + t_i, \quad i = 1, \ldots, m \\
& \boldsymbol{\mu}_0^T\boldsymbol{\phi} - \boldsymbol{\gamma}^T|\boldsymbol{\phi}| \geq \alpha, \\
& \boldsymbol{1}^T\boldsymbol{\phi} = 1 \\
& \left\| \begin{bmatrix} 2\bar{\boldsymbol{D}}^{\frac{1}{2}}\boldsymbol{\phi} \\ 1 - \delta \end{bmatrix} \right\| \leq 1 + \delta
\end{array}$$

With this we are now able to do some computational experiments.

Chapter 5

Calculations

In this Chapter I will present the results of some computational experiments on real market and simulated data. The experiments were designed to compare the classical portfolio approach to the robust reformulation using the different theoretic frameworks introduced in this work.

I will also compare my results to those Goldfarb and Iyengar [2] reported. They always solved the maximum Sharpe Ratio problem to avoid the definition of a certain risk or return level necessary for the minimum variance or the maximum return problem. Therewith they had to face that - especially in difficult markets - the classical maximum Sharpe ratio problem must not have a solution as was described in Chapter 2. Whenever this occurred they abandoned the comparison of the two techniques.

As I retrieved market data between 06.07.1999 - 30.05.2003, when exchange markets were decreasing most of the time, the solution of the classical maximum Sharpe ratio problem became regularly negative. Therefore I solved the minimum variance portfolio problem instead. This has the additional benefit of a realistic scenario. There are quite a few funds around, for example funds underlying capital based life insurances or pension funds, which by law have to be invested in a way that guarantees at least a given minimum percentage on the long run.

To solve the generated SOCP problems I used the Toolbox SeDuMi V1.05 from October 2001 (Toolbox for optimization over self-dual homogeneous cones) [7] within Matlab 5.30 (R11). My hardware was a Compaq Armada M700 Laptop running Windows 2000.

All matlab code '*.m' I wrote to perform these tasks is appended at the end of this work. Moreover, all data sets and program codes which were used for my calculations are on the floppy disc enclosed to this master thesis.

5.1 Calculations on real market data

In the Goldberg study comparisons where made over 10 investment periods with a sample size of 90 per period. This is a reasonable choice. Therefore I used the same number of data points to simplify the comparison of the results. Nevertheless, any existing relevant benefit of the robust

TICKER	Name	Ticker	Name
ADS GY	Adidas-Salomon	LHA GY	Deutsche Lufthansa
AIX GY	Aixtron	LIN GY	Linde
ALT GY	Altana	MAN GY	MAN
ALV GY	Allianz	MDN GY	Medion
AMB2 GY	AMB Generali Holding	MEO GY	Metro
BAS GY	BASF	MGT GY	MG Technologies
BAY GY	Bayer	MLP GY	MLP
BEI GY	Beiersdorf	MOB GY	Mobilcom
BMW GY	Bayerische Motoren Werke	MRK GY	Merck KGaA
BOS3 GY	Hugo Boss	MUV2 GY	Muenchener Rueckversicherungs
BUD GY	Buderus	NDA GY	Norddeutsche Affinerie
BZL GY	Beru	PFV GY	Pfeiffer Vacuum Technology
CBK GY	Commerzbank	PNE2 GY	Plambeck Neue Energien
CLS GY	Celesio	PSM GY	ProSieben SAT.1 Media
CON GY	Continental	PUM GY	Puma Rudolf Dassler Sport
DBK GY	Deutsche Bank	QIA GY	Qien NV
DCX GY	DaimlerChrysler	RHK3 GY	Rhoen Klinikum
DGX GY	Degussa	RHM3 GY	Rheinmetall
DOU GY	Douglas Holding	RWE GY	RWE
DRW3 GY	Draegerwerk	SAP GY	SAP
DTE GY	Deutsche Telekom	SAZ GY	Stada Arzneimittel
DYK3 GY	Dyckerhoff	SCH GY	Schering
EIE3 GY	United Internet	SDF GY	K+S
EOA GY	E.ON	SGL GY	SGL Carbon
FME GY	Fresenius Medical Care	SIE GY	Siemens
FRE3 GY	Fresenius	SMY GY	SCM Microsystems Inc
GBF GY	Bilfinger Berger	SNG GY	Singulus Technologies
HDD GY	Heidelberger Druckmaschinen	SOW4 GY	Software
HEI GY	HeidelbergCement	SRZ GY	Schwarz Pharma
HEN3 GY	Henkel KGaA	SZG GY	Salzgitter
HNR1 GY	Hannover Rueckversicherungs	SZU GY	Suedzucker
HOT GY	Hochtief	TKA GY	ThyssenKrupp
HVM GY	Bayer. Hypo- & Vereinsbank	TPL GY	Teleplan International NV
IDS GY	IDS Scheer	TUI GY	TUI
IKB GY	IKB Deutsche Industriebank	VOS GY	Vossloh
IVG GY	IVG Immobilien	VOW GY	Volkswen
IWK GY	IWKA	WAD3 GY	Wella
JEN GY	Jenoptik	WCM GY	WCM Beteiligungs & Grundbesitz-
KAR GY	KarstadtQuelle	XOS GY	Ixos Software
KRN3 GY	Krones	ZPF GY	Zapf Creation
LEO GY	Leoni		

Table 5.1: set of assets

against the classical approach should be visible for any desired asset opportunity set and any time period. Hence, I retrieved end of day prices of all assets which where members of the HDAX on 30.05.2003 per Open Bloomberg. According to the Open Bloomberg description the 'HDAX Index is a total rate of return index of the 110 most highly capitalized stocks traded on the Frankfurt Stock Exchange.'

Thus, to perform portfolio calculations for ten investment periods I needed 991 prices. This was not available for all of the 110 assets currently in the HDAX. Reducing the opportunity set to those assets with enough data points available, left me with the 81 assets named in Table 5.1.

Just by chance I tried to use the industry indices of the CDAX listed in Table 5.2 as factors for the multi-index model. The data was also retrieved per Open Bloomberg for the same time

TICKER	Name	Ticker	Name
CXPA	Automobiles Index	CXPN	Industrial Index
CXPB	Banks Index	CXPO	Construction Index
CXPC	Chemicals Index	CXPP	Pharmacie & Health Care Index
CXPD	Media Index	CXPR	Retail Index
CXPE	Basic Resources Index	CXPS	Software Index
CXPF	Food & Beverages Index	CXPT	Telecomunication Index
CXPH	Technology Index	CXPU	Utilities Index
CXPI	Insurance Index	CXPV	Financial Services Peformance Index
CXPL	Transport & Logistic Index	CXPY	Consumer Index

Table 5.2: set of subindices

period. The covariance matrix $\boldsymbol{F}$ was calculated separately for each period on the sample data of the previous 90 trading days and was assumed to be fix during each period. Then a regression analysis was run on the data to estimate the vector $\boldsymbol{\mu_0}$ and the factor loading matrix $\boldsymbol{V_0}$ as described in Section 3.4.

To check the quality of the regression I calculated the ratio of the error variance σ_{ϵ_i} to the variance explained by the model $\boldsymbol{V_0^T F V_0}$. The results are given in Table 5.3. Obvious the chosen indices do not very well in explaining the movement of the asset returns.

1	2	3	4	5	6	7	8	9	10	11
1.709	1.543	1.626	2.391	1.204	2.273	1.956	2.153	2.863	2.746	2.908

Table 5.3: error variance / model variance using industry indices

As finding suitable indices is a sophisticated matter and not the topic of this master thesis, I abandoned the real market indices. Instead I used the eigenvectors of the covariance matrix as artificial indices. To get the corresponding returns of these constructed indices, I based them to 1 so that the elements of each vector sum up to 1 and calculated their return according to the sample return for each asset. Again I controlled the quality of the regression by the ratio of the mean error variance σ_{ϵ_i} to the variance explained by the model $\boldsymbol{V_0^T F V_0}$. The results for increasing numbers m of eigenvectors is given in Table 5.4.

m	1	2	3	4	5	6	7	8	9	10	11
10	1.648	1.487	1.509	3.247	1.103	3.681	2.609	2.925	3.060	2.596	2.322
20	0.709	0.597	0.595	1.117	0.518	1.288	1.009	1.196	1.140	1.135	0.961
30	0.409	0.323	0.335	0.551	0.286	0.609	0.539	0.637	0.604	0.553	0.531
40	0.231	0.178	0.181	0.289	0.161	0.321	0.290	0.320	0.309	0.288	0.301

Table 5.4: error variance / model variance

Let us agree to call a relative error variance of 0.3 - 0.2 acceptable small, i.e. the model explains 70% to 80% of the sample variance. Then we see from table 5.4 that we have to consider 40 or more eigenvectors. This seems contrary to the statement of Goldberg and Iyengar [2] that they could reduce the relative error variance of their selection consisting of 43 assets sufficiently by adding only five eigenvectors to their chosen 5 real world indices.

Nevertheless, I accepted the multi-index model with 40 eigenvectors as underlying model for my further calculations. The covariance matrix $\boldsymbol{F}$, the mean return vector $\boldsymbol{\mu_0}$ and the factor loading matrix $\boldsymbol{V_0}$ were extracted from the data as described above. Equation (3.1) was used to calculate the error variance σ_{ϵ_i}. The additional parameters $\boldsymbol{\gamma}, \boldsymbol{\rho}, \boldsymbol{G}$ necessary for the robust approach were calculated according to equations (3.3), (3.6) and (3.8). The upper limit $\bar{d}_i$ was set to σ_{ϵ_i}. Then for each period a classical and a robust portfolio were determined from the $k = 90$ previous sample data points. As the robust approach is dependent on the probability ω I rerun the calculations for the values $\omega = 20\%, 40\%, 60\%, 80\%, 90\%, 95\%, 99\%$. Likewise I used the values 0%, 3%, 5%, 8%, 10%, 12%, 15% for the minimum anualized expected return α.

Considering the high mean of the error term and the relative bad market situation during the sample time periods, it is not surprising, that it was not always possible to find a valid solution for all pairs of ω and α. In table 5.5 every pair for which a solution for all ten investment periods was found, is denoted with a '∘'. Those tuples α, ω were problems occurred are marked with a 'x'.

	$\omega = 20$	$\omega = 40$	$\omega = 60$	$\omega = 80$	$\omega = 90$	$\omega = 95$	$\omega = 99$
$\alpha =$ 0 %	∘	∘	∘	∘	∘	∘	x
$\alpha =$ 3 %	∘	∘	∘	∘	∘	x	x
$\alpha =$ 5 %	∘	∘	∘	∘	∘	x	x
$\alpha =$ 8 %	∘	∘	∘	∘	∘	x	x
$\alpha =$ 10 %	∘	∘	∘	∘	x	x	x
$\alpha =$ 12 %	∘	∘	∘	∘	x	x	x
$\alpha =$ 15 %	∘	∘	∘	∘	x	x	x

Table 5.5: sucessful calculations on real market data

For those α, ω-pairs where solutions existed I compared the performance of the classical portfolio $\boldsymbol{\phi_c}$ and the robust portfolio $\boldsymbol{\phi_r}$ with respect to the following aspects:

5.2 Performance measures used for comparison

For both portfolios $\boldsymbol{\phi_c}$ and $\boldsymbol{\phi_r}$ the **mean Sharpe ratio** given by

$$\frac{(\boldsymbol{\mu_0} - r_f\mathbf{1})^T\boldsymbol{\phi}}{\sqrt{\boldsymbol{\phi^T}(\boldsymbol{V_0^T F_p V_0})\boldsymbol{\phi}}}$$

was calculated with the risk free rate $r_f = 0$. Because the portfolio valid in period $(p+1)$ is determined by an optimization on the data of period p, while the mean Sharpe Ratio of period $(p+1)$ is calculated on the market data of period $(p+1)$, it is possible that the Sharpe ratio of one or both portfolios ϕ_r, ϕ_c can become negative. So to compare the fraction of the two values is no use. Thus I plotted the results separately next to each other (see Appendix A).

Analogous plots - also with $r_f = 0$ - were created for the **worst case Sharpe ratio**, which is the solution of the following minimization problem:

$$\min_{\{\boldsymbol{V}\in S_v, \boldsymbol{\mu}\in S_\mu, \boldsymbol{D}\in S_d\}} \frac{(\boldsymbol{\mu} - r_f\mathbf{1})^T\boldsymbol{\phi}}{\sqrt{\boldsymbol{\phi^T}(\boldsymbol{V^T F V})\boldsymbol{\phi}}}$$

Again in investment period $p+1$ the portfolios calculated from the data of period p were measured on the data of period $p+1$.

Another interesting feature is the **wealth** w of each portfolio at time $t+1$ and can by calculated by:

$$\begin{aligned} w_r^{(t+1)} &= \left[\left(\prod_{tp\le k\le (t+1)p} (1+r_k)^T\right)\phi_r^t\right] w_r^t \\ w_m^{(t+1)} &= \left[\left(\prod_{tp\le k\le (t+1)p} (1+r_k)^T\right)\phi_m^t\right] w_m^t \end{aligned}$$

Here I plotted the relative wealth.

Last but not least I compared the **transaction costs** according to the following equation:

$$||\boldsymbol{\phi_r^t} - \boldsymbol{\phi_r^{(t+1)}}||_1$$

The robust transaction costs is always plotted relative to the classical transaction costs.

5.3 Results of the calculations on real market data

The plots of the results on the real market data can be looked up in Appendix A. The different plots do not vary much for the different values of α. So I will concentrate on the comparison of the four performance measures described in the previous section.

5.3.1 Mean Sharpe ratio

For the given real market data the mean Sharpe ratio of the robust approach increases in most periods with ω - although not in all. In period 6 the mean Sharpe ratio for $\omega = 90$ is even the lowest. At the significance level of 90% and in most periods already for $\omega = 80\%$ the robust mean Sharpe ratio is higher than the mean Sharpe ratio of the classical portfolio. In periods 2 and 4 the mean Sharpe ratio of the robust strategy is even positive for all ω while the ratio of the classical approach is negative.

Table 5.6 denotes with a '+' if the robust mean Sharpe ratio is better, '-' if it is worse and '$\sim$' if it is similar to the classical mean Sharpe ratio based on the data for $\alpha = 5\%$.

ω	1	2	3	4	5	6	7	8	9	10
20	+	+	+	+	+	+	-	+	-	+
40	+	+	+	+	-	+	$\sim$	+	$\sim$	+
60	+	+	+	+	-	+	+	+	+	+
80	+	+	+	+	-	+	+	+	+	+
90	+	+	+	+	+	+	+	+	+	+

Table 5.6: comparison mean Sharpe ratios on real market data

5.3.2 Worst case Sharpe ratio

The worst case Sharpe ratio of the robust portfolio is for each ω in almost all periods better than that of the classical portfolio. For $\omega = 90\%$ it outperforms the classical portfolio in every period by far.

Nevertheless, it is important to note that the worst case Sharpe ratios for different ω are not really comparable, because the confidence level ω is used not only for the determination of the portfolio but also for the calculation of the worst case Sharpe ratio. This means that a portfolio with a lower ω is not only created on a smaller uncertainty set, but is also tested on a less worse scenario than a portfolio with a higher ω. It might be interesting to use different values of ω only during the calculation of the portfolios but to test them all on a 95% worst case scenario based on the data of the following period.

Table 5.7 denotes with a '+' if the robust mean Sharpe ratio is better, '-' if it is worse and '$\sim$' if it is similar to the classical mean Sharpe ratio based on the data for $\alpha = 5\%$.

5.3.3 Relative transaction costs

For the given data the transaction costs of the robust portfolio are in general higher than that of the classical strategy and increase with ω.

ω	1	2	3	4	5	6	7	8	9	10
20	+	+	+	+	+	+	+	+	-	+
40	+	+	+	+	-	+	+	+	~	+
60	+	+	+	+	-	+	+	+	+	+
80	+	+	+	+	~	+	+	+	+	+
90	+	+	+	+	+	+	+	+	+	+

Table 5.7: comparison worst Sharpe ratios on real market data

5.3.4 Relative wealth

For $\omega \geq 80\%$ the robust strategy outperforms the classical by far. Here the result is dependent on α. For small α - 0% and 3% - the robust approach is about 60% to 70% better than the classical. For medium α - 5% and 8% - the outperformance is more than 70% and for large α - 10% and 12% - it's just about 35% to 45% for the significance level $\omega = 90\%$.

5.3.5 Absolute values

The absolute values for $\alpha = 5\%$ and $\omega = 90\%$ are given in Table 5.8 where R denotes the robust and C the classical portfolio:

value	1	2	3	4	5	6	7	8	9	10	11
σ_R^2	6.026	5.498	3.711	2.018	7.951	1.650	4.001	2.467	1.781	3.451	1.173
costs_R		1.871	1.612	1.824	1.992	1.986	1.551	1.733	1.688	1.730	
wealth_R	1.000	1.179	1.533	1.338	1.479	1.542	1.663	1.785	1.898	1.835	2.099
σ_C^2	0.302	0.734	0.621	0.140	0.871	0.119	0.268	0.176	0.162	0.222	0.127
costs_C		1.368	1.230	1.253	1.144	1.246	0.854	1.212	1.149	0.955	
wealth_C	1.000	1.070	1.335	1.187	1.216	1.162	1.161	1.171	1.146	1.140	1.200

Table 5.8: absolute values on real market data

5.4 Calculations on simulated data

We have seen from Table 5.4 that the chosen market data does not covary very strongly. This made it impossible to find valid solutions for high values of ω. Thus I simulated three data sets again containing 990 samples for 81 assets following a browning motion with constant drift and volatility. With each simulation I increased the covariance of the assets. In table 5.9 the mean error variance is given relative to the model variance.

sim	1	2	3	4	5	6	7	8	9	10	11
sim1	0.129	0.114	0.102	0.126	0.080	0.128	0.148	0.103	0.130	0.109	0.145
sim2	0.065	0.057	0.049	0.054	0.067	0.059	0.058	0.049	0.069	0.082	0.061
sim3	0.003	0.003	0.002	0.003	0.002	0.002	0.002	0.003	0.002	0.002	0.003

Table 5.9: mean error variance / model variance of simulated data sets

Again it was not possible to determine solutions for all pairs of ω, α. Tables 5.10 to 5.12 represent each pair where a solution was found for all periods with a '∘' and those where problems occurred with 'x'.

	$\omega = 20$	$\omega = 40$	$\omega = 60$	$\omega = 80$	$\omega = 90$	$\omega = 95$	$\omega = 99$
$\alpha =$ 0 %	∘	∘	∘	∘	∘	∘	x
$\alpha =$ 3 %	∘	∘	∘	∘	∘	x	x
$\alpha =$ 5 %	∘	∘	∘	∘	∘	∘	x
$\alpha =$ 8 %	∘	∘	∘	∘	∘	∘	x
$\alpha =$ 10 %	∘	∘	∘	∘	∘	x	x
$\alpha =$ 12 %	∘	∘	∘	∘	∘	x	x
$\alpha =$ 15 %	∘	∘	∘	∘	∘	x	x

Table 5.10: successful calculations of simulation 1

	$\omega = 20$	$\omega = 40$	$\omega = 60$	$\omega = 80$	$\omega = 90$	$\omega = 95$	$\omega = 99$
$\alpha =$ 0 %	∘	∘	∘	∘	∘	∘	∘
$\alpha =$ 3 %	∘	∘	∘	∘	∘	x	x
$\alpha =$ 5 %	∘	∘	∘	∘	∘	∘	x
$\alpha =$ 8 %	∘	∘	∘	∘	∘	x	x
$\alpha =$ 10 %	∘	∘	∘	∘	∘	x	x
$\alpha =$ 12 %	∘	∘	∘	∘	∘	x	x
$\alpha =$ 15 %	∘	∘	∘	∘	∘	x	x

Table 5.11: successful calculations of simulation 2

	$\omega = 20$	$\omega = 40$	$\omega = 60$	$\omega = 80$	$\omega = 90$	$\omega = 95$	$\omega = 99$
$\alpha =$ 0 %	∘	∘	∘	∘	∘	∘	∘
$\alpha =$ 3 %	∘	∘	∘	∘	∘	∘	∘
$\alpha =$ 5 %	∘	∘	∘	∘	∘	∘	∘
$\alpha =$ 8 %	∘	∘	∘	∘	∘	∘	∘
$\alpha =$ 10 %	∘	∘	∘	∘	∘	∘	∘
$\alpha =$ 12 %	∘	∘	∘	∘	∘	∘	∘
$\alpha =$ 15 %	∘	∘	∘	∘	∘	∘	∘

Table 5.12: successful calculations of simulation 3

For those pairs of ω and α were a solution existed I used the same performance measures as specified on the real market data in section 5.2. The plotted results of each simulation are in the Appendices B - C. Again, it can be seen that the results of any of the analyzed values does not vary much with α on all simulated data sets. In the following three sections I will compare the robust and classical solutions separately for each simulation on their mean Sharpe ratio, worst Sharpe ratio, transaction costs and the relative wealth. Further I will have a short look on the absolute wealth of the robust portfolio dependent on ω and the absolute results of each solution.

5.5 Results of simulation 1

5.5.1 Mean Sharpe ratio

On this data set it can be clearly seen that the mean Sharpe ratio of the robust approach does not always increase with ω. In period 8 the mean Sharpe ratio for $\omega = 95\%$ is even the worst of all. But on this data set it is still true for most periods that the larger ω the better the mean Sharpe ratio of the robust portfolio.

Nevertheless, is the mean Sharpe ratio of the robust portfolio in all periods except of period 9 notedly greater than that of the classical portfolio if $\omega = 90\%$ or $\omega = 95\%$ and in most periods already for $\omega = 80\%$. Again there is a period - period 1 - were the robust mean Sharpe ratio is even positive for $\omega \geq 80\%$ while the classical is negative.

Table 5.13 denotes with a '+' if the robust worst Sharpe ratio is better, '-' if it is worse and '$\sim$' if it is similar to the classical worst Sharpe ratio.

ω	1	2	3	4	5	6	7	8	9	10
20	$\sim$	-	$\sim$	-	$\sim$	+	+	+	+	+
40	+	-	$\sim$	-	+	$\sim$	+	+	+	+
60	+	-	$\sim$	-	+	+	+	+	-	+
80	+	$\sim$	+	-	+	+	+	+	-	+
90	+	+	+	$\sim$	+	+	+	+	-	+
95	+	+	+	+	+	+	+	+	+	+

Table 5.13: robust mean Sharpe ratio / classical mean Sharpe ratio simulation 1

5.5.2 Worst case Sharpe ratio

For $\omega \geq 95\%$ the worst case Sharpe ratio of the robust approach is greater than that of the classical portfolio for all periods.

Table 5.14 denotes with a '+' if the robust worst Sharpe ratio is better, '-' if it is worse than and '$\sim$' if it is similar to the classical worst Sharpe ratio for $\alpha = 5\%$:

ω	1	2	3	4	5	6	7	8	9	10
20	$\sim$	-	$\sim$	-	$\sim$	+	+	+	+	+
40	+	-	+	-	+	$\sim$	+	+	+	+
60	+	-	+	-	+	+	+	+	-	+
80	+	$\sim$	+	$\sim$	+	+	+	+	-	+
90	+	+	+	+	+	+	+	+	-	+
95	+	+	+	+	+	+	+	+	+	+

Table 5.14: robust worst Sharpe ratio / classical worst Sharpe ratio simulation 1

5.5.3 Relative transaction costs

For the data set of the first simulation the transaction costs of the robust strategy are sometimes higher and sometimes lower than that of the classical strategy. A relation between ω and the transaction costs cannot be observed.

5.5.4 Wealth

For $\omega \geq 80\%$ the robust strategy outperforms the classical by far. Here a dependence on α is not obvious. The absolute wealth of the robust portfolio increases with ω.

5.5.5 Absolute values of simulation 1

The absolute values for $\alpha = 5\%$ and $\omega = 90\%$ are given in Table 5.15 where R denotes the robust and C the classical portfolio:

value	1	2	3	4	5	6	7	8	9	10	11
σ^2 R	52.50	64.62	64.91	61.40	86.27	56.70	49.33	69.68	58.98	60.28	50.36
Costs R		1.883	1.837	2.000	1.859	2.000	2.000	2.000	1.707	1.982	
Wealth R	1.000	0.424	0.801	0.705	0.922	0.888	0.716	2.091	1.167	1.563	1.460
σ^2 C	37.98	44.31	48.23	42.03	65.60	37.40	33.41	51.09	38.18	45.57	34.82
Costs C		2.000	2.000	1.975	1.976	1.943	1.955	1.886	1.833	1.884	
Wealth C	1.000	0.330	0.561	0.486	0.645	0.539	0.398	1.099	0.525	0.747	0.606

Table 5.15: absolute values on data set of simulation 1

5.6 Results of simulation 2

5.6.1 Mean Sharpe ratio

Obviously for this simulation the previous observation that in general the larger ω the better the mean Sharpe ratio of the robust portfolio does not hold. See for example the periods 2,3,4,5,6 or 7. In period 7 the mean Sharpe ratio even is worst for $\omega = 90\%$. Here the mean Sharpe ratio of the robust approach is in some periods notedly greater than and in some notedly less than the mean Sharpe ratio of the classical portfolio for all significance levels ω. In period 10 it can even be observed that the robust ratio is negative for $\omega = 90\%$ while the classical is positive. Table 5.16 presents the '+','-' comparison of the two portfolios for $\alpha = 0\%$:

5.6.2 Worst case Sharpe ratio

For this simulation the worst case Sharpe ratio of the robust approach is not always better than that of the classical for any ω. Table 5.17 contains the '+','-' comparison of the two portfolios for $\alpha = 0\%$.

ω	1	2	3	4	5	6	7	8	9	10
20	+	-	+	-	-	+	-	+	+	+
40	+	-	+	-	+	+	-	+	+	~
60	+	-	~	-	+	+	-	~	+	-
80	+	-	+	-	+	+	-	+	+	-
90	+	-	+	-	+	+	-	+	+	-
95	+	-	+	-	-	+	-	+	+	-
99	+	-	+	+	-	+	-	+	+	-

Table 5.16: robust mean Sharpe ratio / classical mean Sharpe ratio simulation 2

ω	1	2	3	4	5	6	7	8	9	10
20	+	-	+	-	-	+	-	+	+	+
40	+	-	+	-	+	+	-	+	+	+
60	+	-	+	-	+	+	-	+	+	-
80	+	-	+	-	+	+	-	+	+	-
90	+	-	+	-	+	+	-	+	+	-

Table 5.17: robust worst Sharpe ratio / classical worst Sharpe ratio simulation 2

5.6.3 Relative transaction costs

In this simulation we find again that the transaction costs of the robust strategy are for some periods higher, but also for some periods lower than the costs of the classical strategy. Here the transaction costs are less the higher ω. But in some periods the transaction costs are even the same for all ω.

5.6.4 Relative wealth

For $\omega = 80\%$ the robust strategy outperforms the classical by far in period 9 and 10 but underperforms the classical in period 5. For $\omega = 90\%$ the robust strategy outperforms the classical by far in period 9,10 and 11 but underperforms the classical in periods 3,5-8.

5.6.5 Absolute values of simulation 2

The absolute values for $\alpha = 5\%$ and $\omega = 90\%$ are given in Table 5.18 where R denotes the robust and C the classical portfolio:

value	1	2	3	4	5	6	7	8	9	10	11
σ^2 R	101.6	120.5	135.6	134.0	102.1	112.4	108.6	137.5	94.33	88.09	114.0
Costs R		2.000	1.395	2.000	2.000	2.000	2.000	2.000	1.774	2.000	
Wealth R	1.000	0.391	0.240	0.181	0.025	0.096	0.621	0.239	0.276	0.074	0.039
σ^2 C	79.285	96.798	107.9	103.8	76.79	89.90	89.00	104.2	74.34	64.01	89.20
Costs C		2.000	2.000	2.000	2.000	1.936	2.000	2.000	2.000	2.000	
Wealth C	1.000	0.357	0.246	0.182	0.027	0.103	0.656	0.302	0.231	0.059	0.040

Table 5.18: absolute values on the data set of simulation 2

5.7 Results of simulation 3

5.7.1 Mean Sharpe ratio

Again the mean Sharpe ratio of the robust approach is not a linear function of ω. See for example period 5 where the mean Sharpe ratio is worst for $\omega = 99\%$. But here in most periods we observe that the larger omega the better the mean Sharpe ratio of the robust portfolio. In periods 7,8,10 the robust mean Sharpe ratio is even positive for $\omega \geq 90\%$ while the classical is negative. For $\omega ge 95\%$ the robust mean Sharpe ratio is better than the classical, except for period 9. Table 5.19 represents the '+','-' comparison of the two portfolios for $\alpha = 0\%$:

ω	1	2	3	4	5	6	7	8	9	10
20	+	-	-	-	~	+	~	+	+	+
40	+	-	-	-	~	+	~	+	+	+
60	+	-	-	-	-	+	~	+	+	+
80	+	-	-	-	-	+	~	+	+	+
90	+	+	-	-	-	+	+	+	+	+
95	+	+	-	+	-	+	+	+	+	+
99	+	+	+	+	-	+	+	+	+	+

Table 5.19: robust mean Sharpe ratio / classical mean Sharpe ratio simulation 3

5.7.2 Worst case Sharpe ratio

The worst case Sharpe ratio of the robust approach is for $\omega = 99\%$ for all periods - except period 5 - better than the classical. Following is the '+','-' comparison of the two portfolios for $\alpha = 0$: controlled for $\alpha = 0,5$

ω	1	2	3	4	5	6	7	8	9	10
20	+	-	-	-	~	+	~	+	+	+
40	+	-	-	-	~	+	~	+	+	+
60	+	-	-	-	-	+	~	+	+	+
80	+	~	-	-	-	+	~	+	+	+
90	+	+	-	-	-	+	+	+	+	+
95	+	-	-	+	-	+	+	+	+	+
99	+	+	~	+	-	+	+	+	+	+

Table 5.20: robust worst Sharpe ratio / classical worst Sharpe ratio simulation 3

5.7.3 Relative transaction costs

Here the transaction costs of the robust portfolio are sometimes less than that of the classical strategy and the higher ω the less the transaction costs.

5.7.4 Relative wealth

For $\omega \geq 90\%$ the robust strategy outperforms the classical by far and the terminal relative wealth increases with α. The greater ω the greater the absolute wealth of the robust portfolio.

5.7.5 Absolute values of simulation 3

The absolute values for $\alpha = 5$ and $\omega = 90$ are given in the Table 5.21 where R denotes the robust and C the classical portfolio:

value	1	2	3	4	5	6	7	8	9	10	11
σ^2 R	85.29	97.32	115.2	97.58	111.7	110.9	117.6	76.54	108.7	112.5	90.80
Costs R		2.000	1.828	2.000	2.000	2.000	1.959	2.000	2.000	2.000	
Wealth R	1.000	1.188	1.442	1.374	0.927	2.445	2.950	0.480	0.494	0.489	0.346
σ^2 C	81.39	92.66	110.6	93.51	105.9	105.6	112.8	72.18	104.6	107.4	85.78
Costs C		2.000	2.000	2.000	2.000	2.000	2.000	2.000	2.000	2.000	
Wealth C	1.000	1.121	1.342	1.339	0.902	2.390	2.663	0.414	0.408	0.389	0.243

Table 5.21: absolute values on the data set of simulation 3

Chapter 6

Conclusions

In the previous chapter I have described the results I found on the real market data and the three simulated data sets. The aim of this chapter is to summarize the results and compare them with the report of Goldfarb and Iyengar [2]. Hence this chapter contains four sections responding to the analyzed performance measures: Sharpe ratio, transaction costs, wealth and absolute values. The fifth section completes this master thesis with a interpretation of the results and a short outlook on possible further research of the topic.

6.1 Mean and worst case Sharpe ratio

For all data sets it was found that the mean Sharpe ratio of the robust strategy does not automatically increase with ω. However, on the real market data and the simulations 1 and 3 it was observed that on the whole the higher ω the better the robust mean Sharpe ratio. On these data sets the mean Sharpe ratio of the robust approach also outperforms in general the classical strategy for large $\omega \geq 90\%$. There were even periods where the robust appoach had a positive return while the classical portfolio faced losses.

The only exception was the data set of the second simulation. Here the mean Sharpe ratio of the robust approach did get worse with increasing ω and the robust portfolio underperformed the classical strategy in several periods - though not in all - even for high $\omega \geq 90$.

This is contrary to the results Goldfarb and Iyengar [2] report. They observe an decrease in the mean Sharpe ratio of the robust portfolio relative to the classical. The reason for this may be that he probably measured his portfolios on the same data sets they were calculated from. Nevertheless, if we are interested in an investment strategy we must measure the portfolio on the sample data during the period when it is valid and not on the sample data on which it is constructed.

The worst case Sharpe ratio of the robust approach was in most periods clearly superior to the classical strategy. Surprisingly this was the case for all values of ω. On the simulated data sets 1 and 3 the worst case Sharpe ratio of the robust was on overall better than the ratio of the classical portfolio for large values of $\omega \geq 90\%$. The Exception is again the data set of simulation 2 where sometimes the robust and sometimes the classical portfolio do better for any value of ω.

With respect to the worst case Sharpe ratio Goldfarb and Iyengar [2] present similar results. They even state that the robust approach does always better than the classical, but they probably calculated the measures against the data period on which the portfolios were determined.

6.2 Transaction costs

Goldfarb and Iyengar [2] state that transaction costs are not only dependent on ω but decrease with high $\omega \geq 90\%$ and are then significantly lower than that of the classical investment strategy. In contrast to that I found that ransaction costs behaved different on each data set. While on the real market data the transaction costs of the robust strategy were higher this could not be found on the simulated data. Here the transaction costs were sometimes higher and sometimes lower than those of the classical strategy. The transaction costs differ for different values of ω, but there was no rule observable how transaction cost vary with ω.

6.3 Wealth

The calculations on the wealth of the portfolio confirm Goldfarb and Iyengard's [2] results. Their statement that the wealth depends strongly on ω proofed true on all four data sets. While the robust approach outperformed the classical portfolio toweringly for the real market data set and simulations 1 and 3 for $omega \geq 80\%$, this cannot be guaranteed for any data set. In simulation 2 it did sometimes worse than the classical portfolio. But whenever it outperformed the classical approach the difference was much higher than when it did underperform.

6.4 Absolute values

The risk of the robust portfolio is always higher than that of the classical portfolio. This is to be expected because both portfolios aim to guarantee a certain level of return, but the robust approach considers also bad market movements thus must take a higher risk to guarantee the prescribed level. The difference in the variances of the robust to the classical portfolio is decreasing with increasing covariance of the asset returns.

The robust portfolio is earning money on market data and sim 1 and loosing money on sim 2 and sim 3. The classical portfolio is also facing losses on the simulation 1.

6.5 Interpretation and Outlook

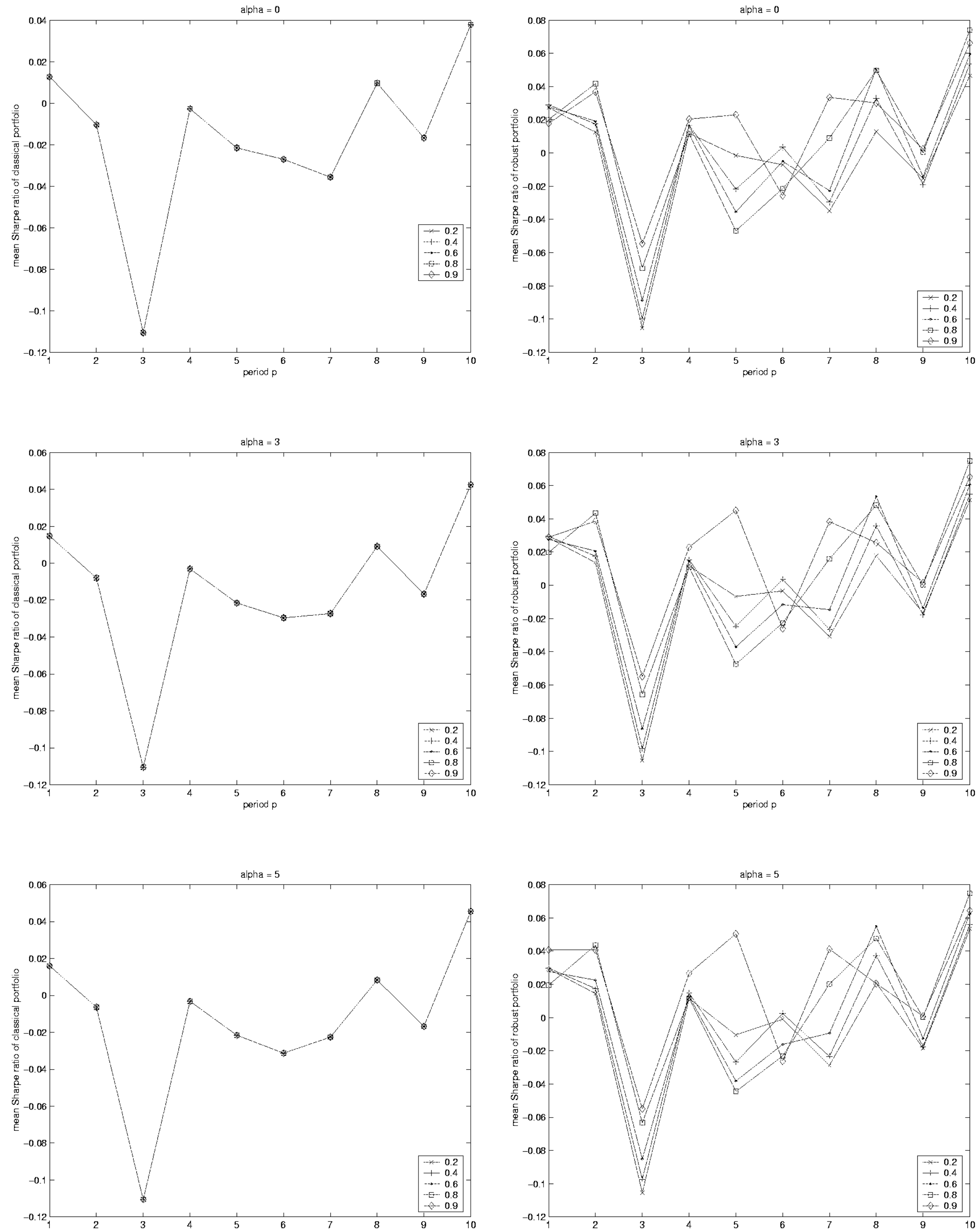
The results on the four data sets let us conclude that the robust portfolio strategy is certainly interesting since it seem to improve the portfolio performance in most situations. Nevertheless, simulation 2 contains the warning that it is also possible to do even worse by investing according to the robust strategy. A likely reason for this observation is, that we estimated our parameters on historical data. Thus we assumed that the data will behave in the same manner in the future than it did in the recent history. The uncertainty set of the robust calculation reflects the historic variance. If this does not reflect the future variance but underestimates the future variance of some assets which are going to loose money and overestimates the variance of future winners, the robust portfolio may be optimized in the wrong direction and does even worse than the classical portfolio in the next period.

Therefore to get some deeper insight in the benefits and risks of the robust approach further research should consider the behaviour of the market data. If the thought formulated above is correct, further improvement should be possible by using advanced estimators of future returns and volatilities, e.g. ARIMA models. It is to be expected that the better the uncertainty sets forecast future risk and returns the more will the robust approach triumph over the classical strategy.

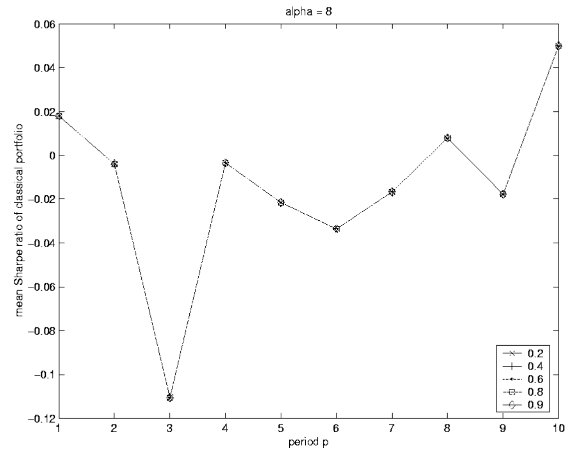
Appendix A

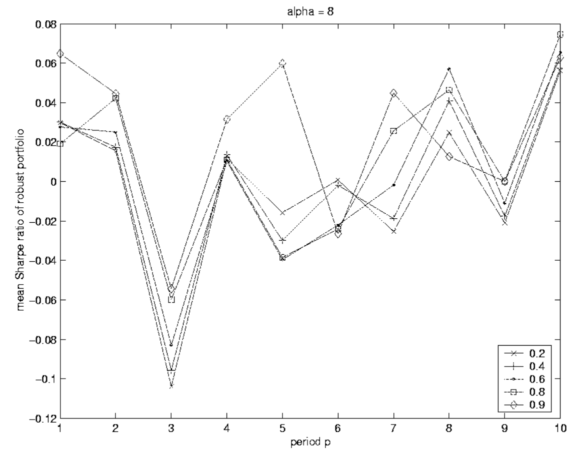
Graphical Results - Market Data

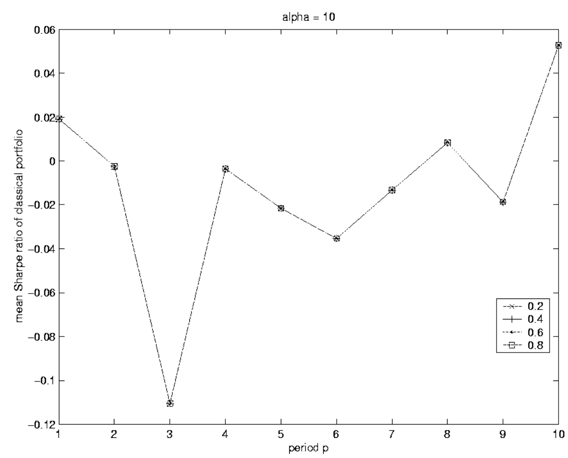
Market Data:
Mean Sharpe Ratio

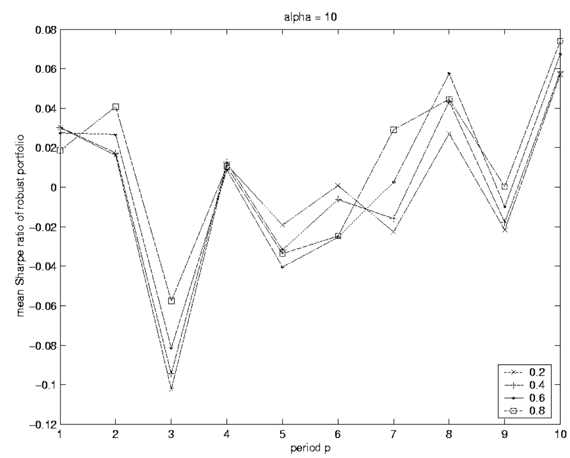


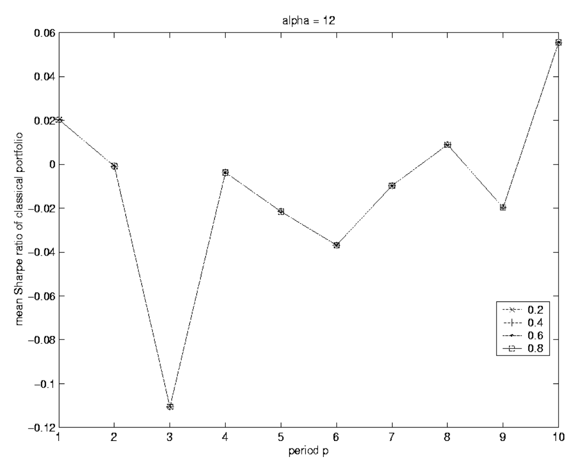
Market Data:
Mean Sharpe Ratio

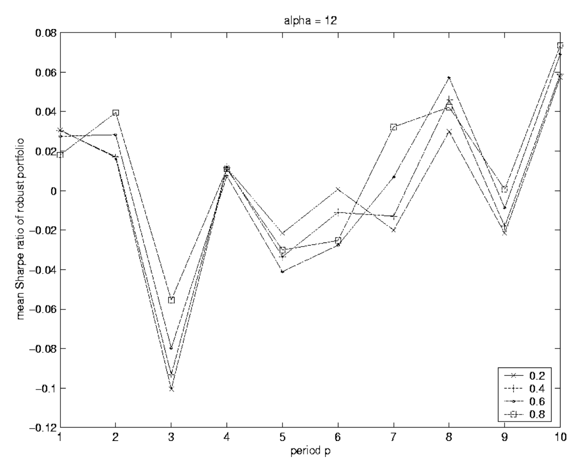




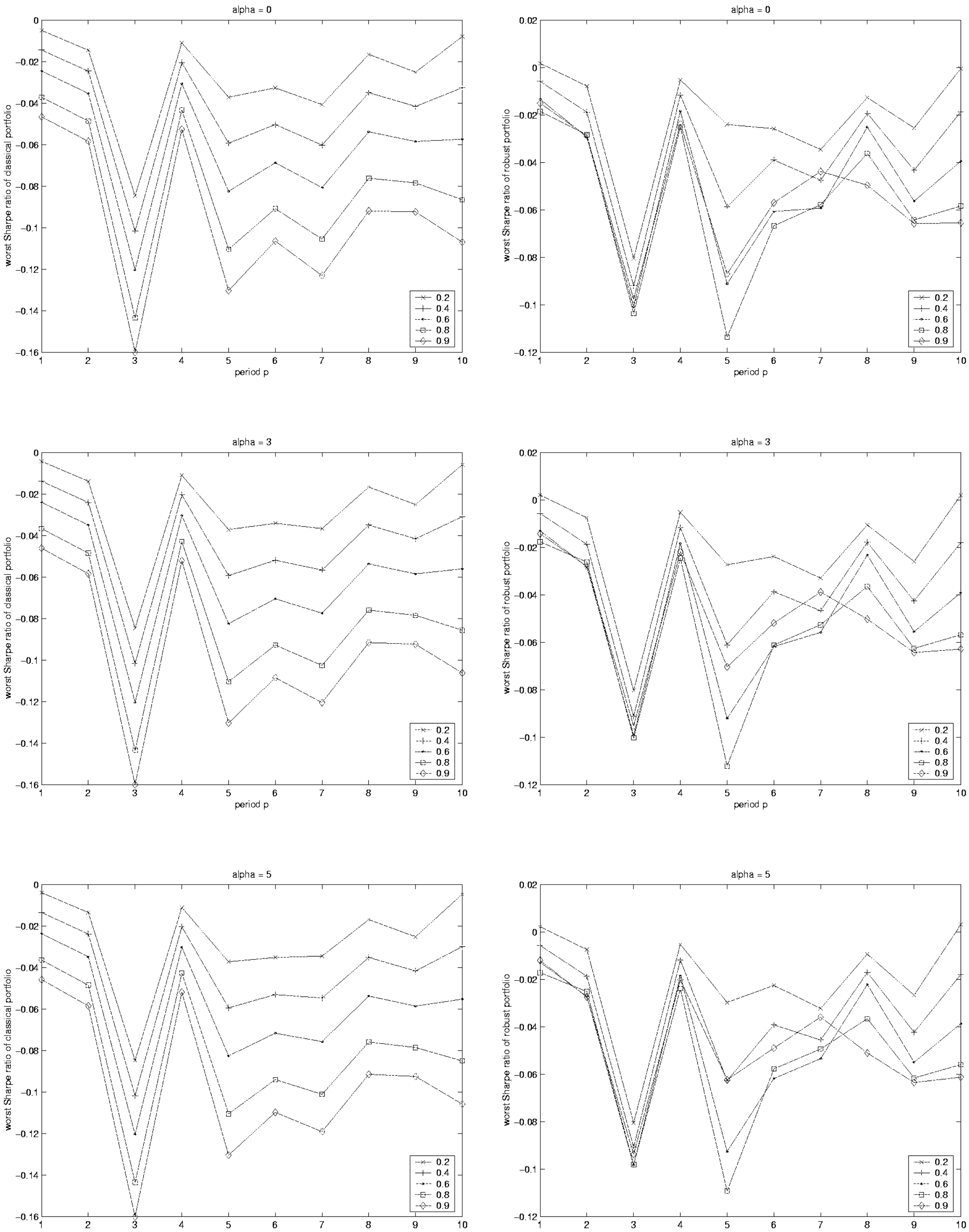




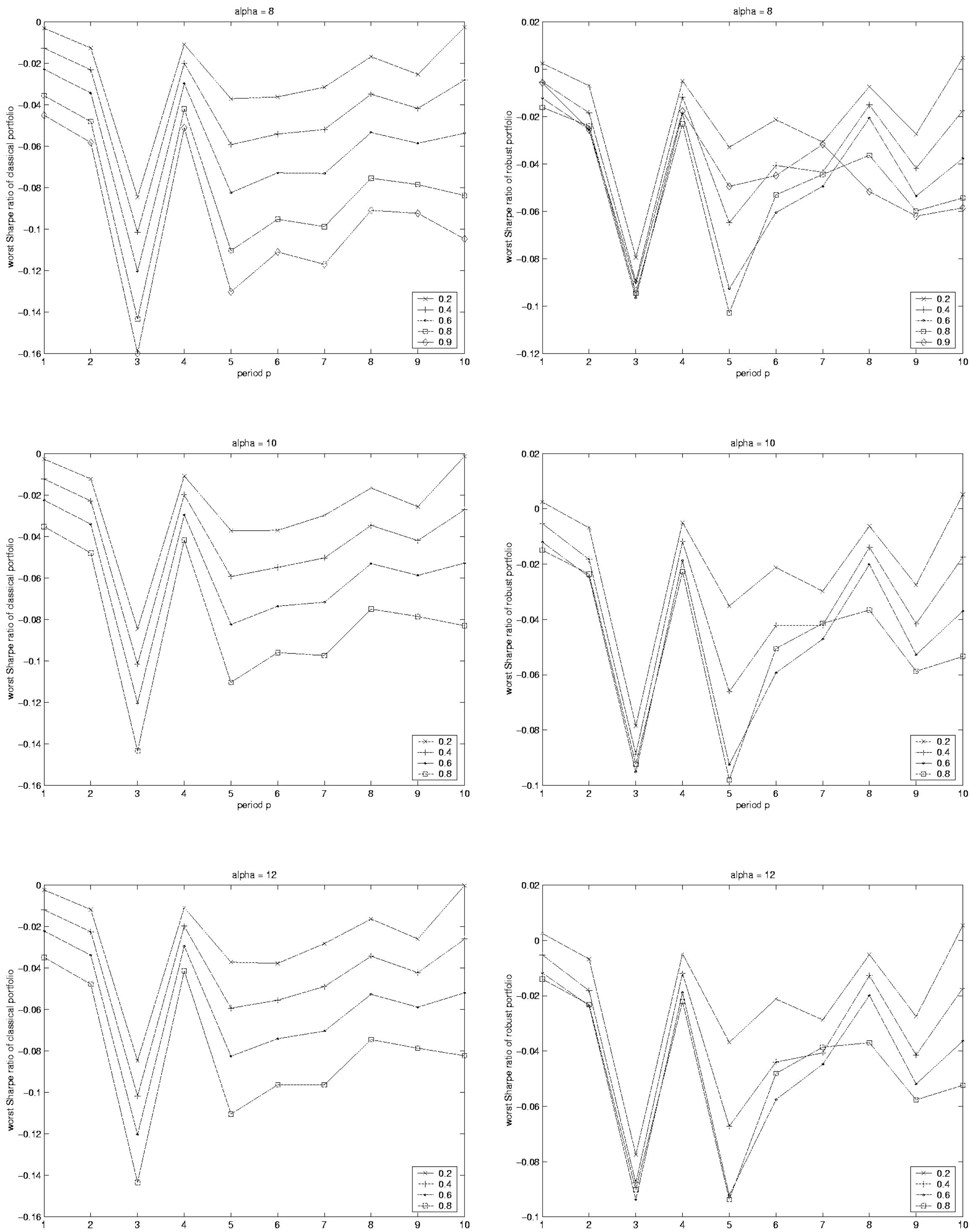




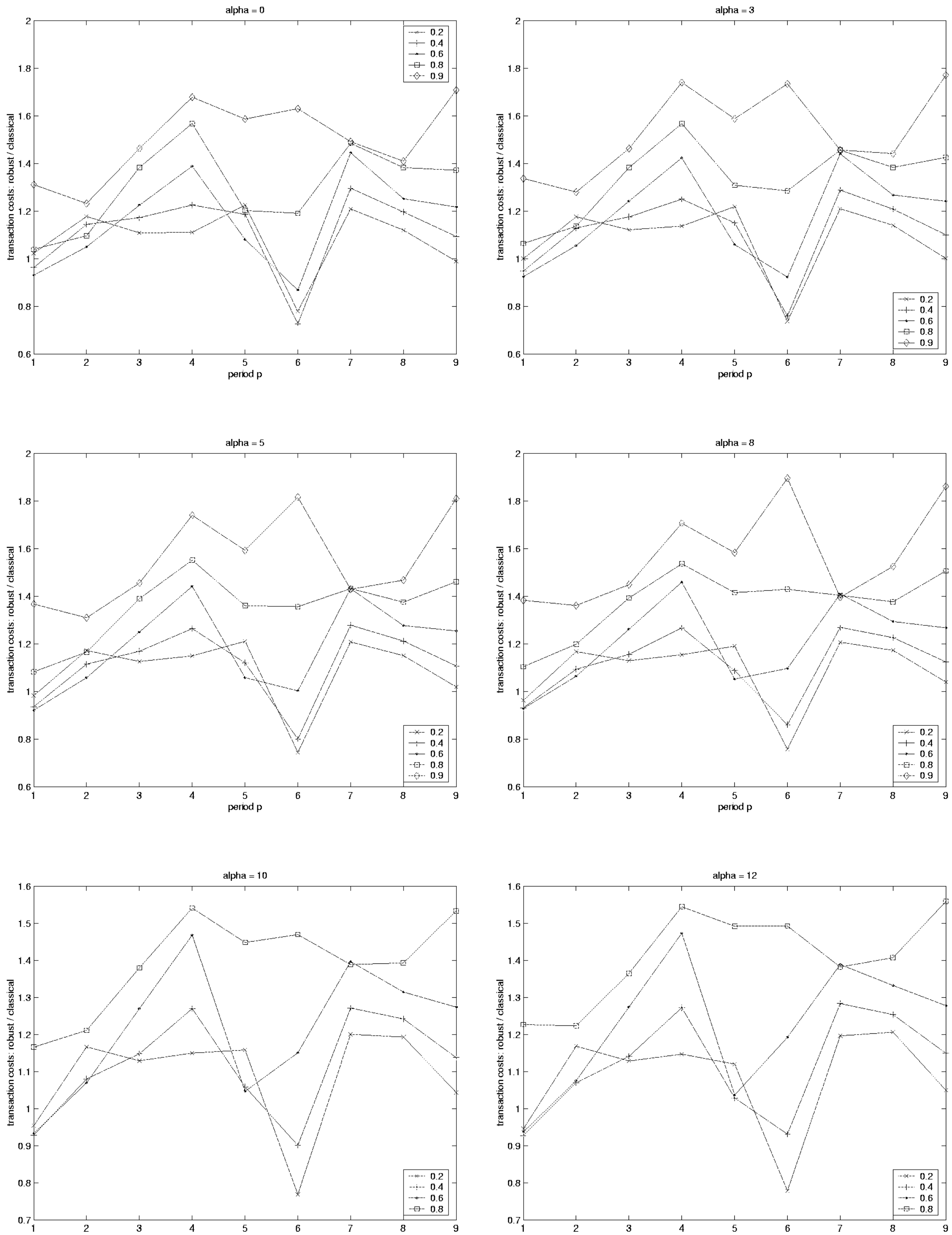
Market Data:
Worst Case Sharpe Ratio



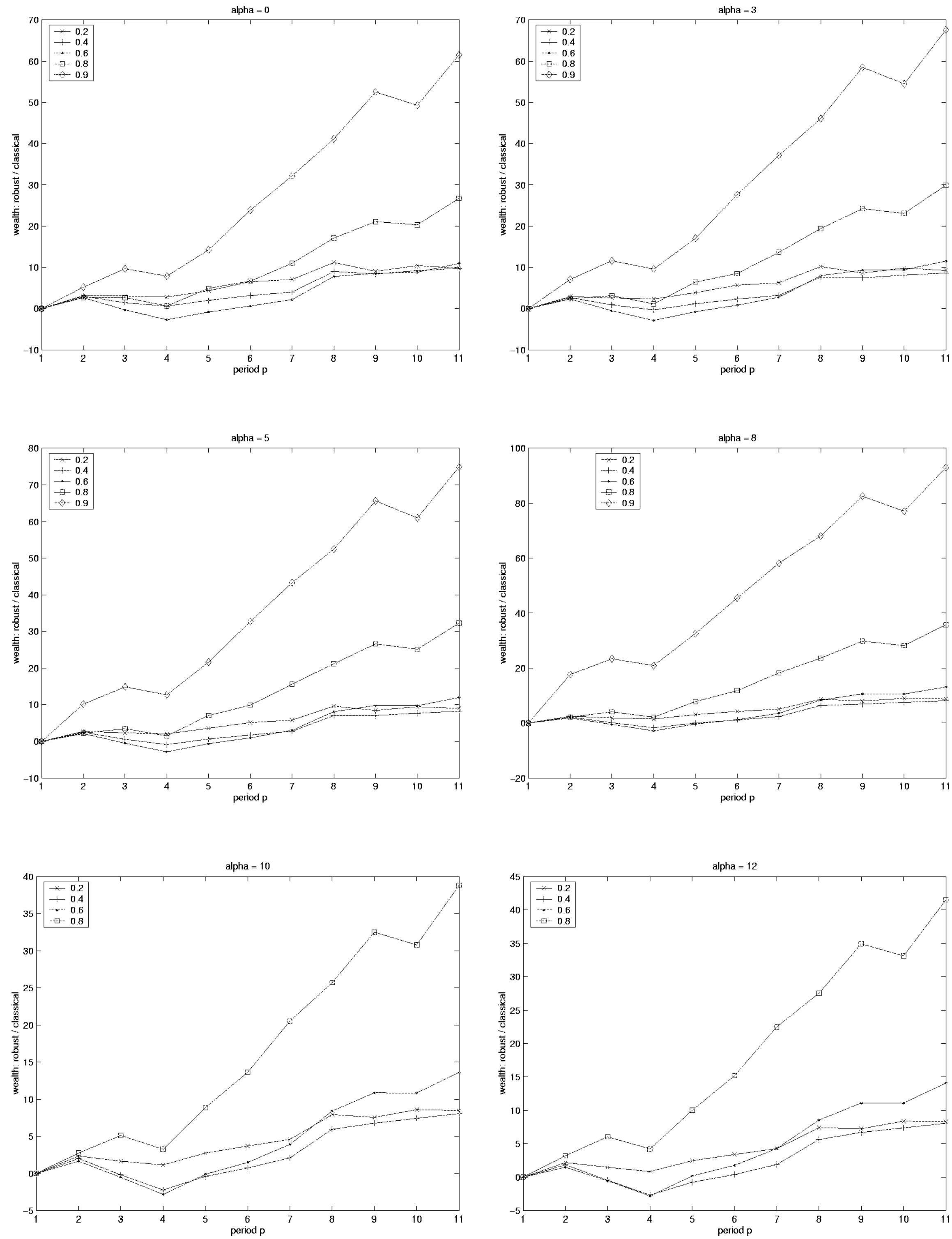
Market Data:
Worst Case Sharpe Ratio



Market Data: Relative Transaction Costs



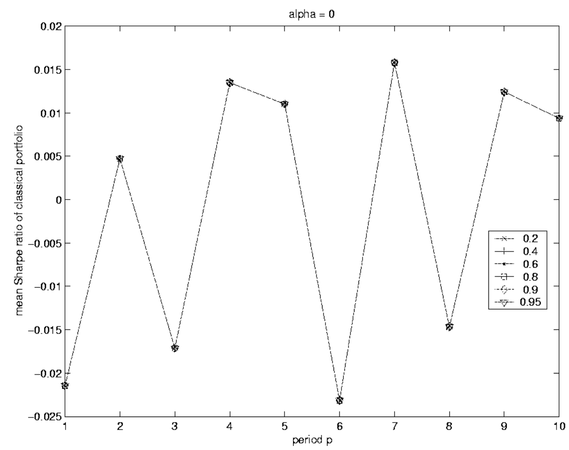
Market Data:
Relative Wealth

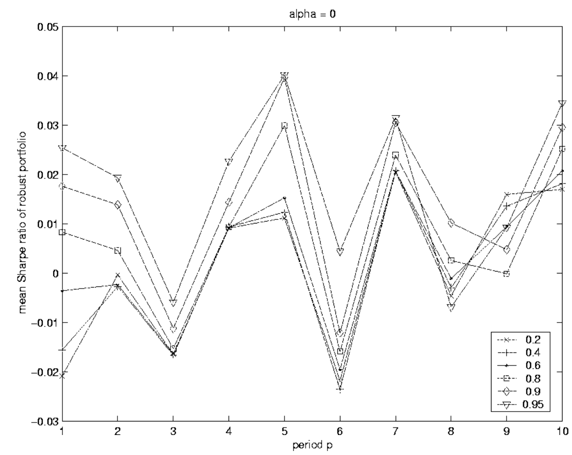


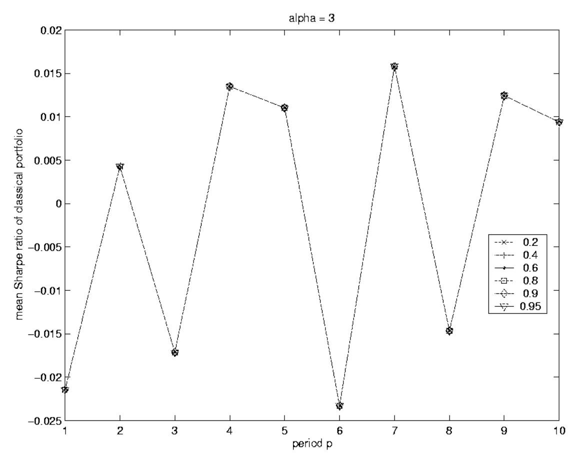
Appendix B

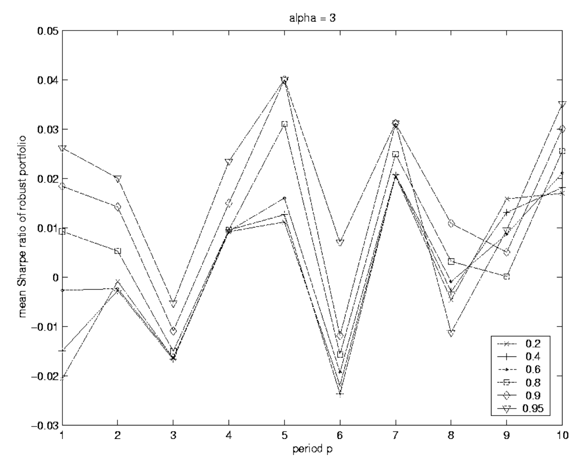
Graphical Results - Simulation 1

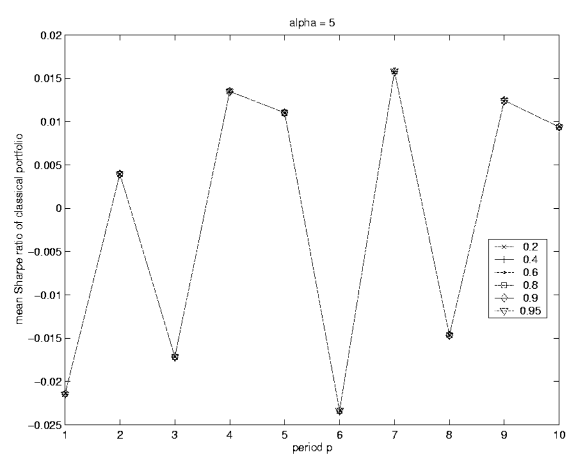
Data of Simulation 1
Mean Sharpe Ratio

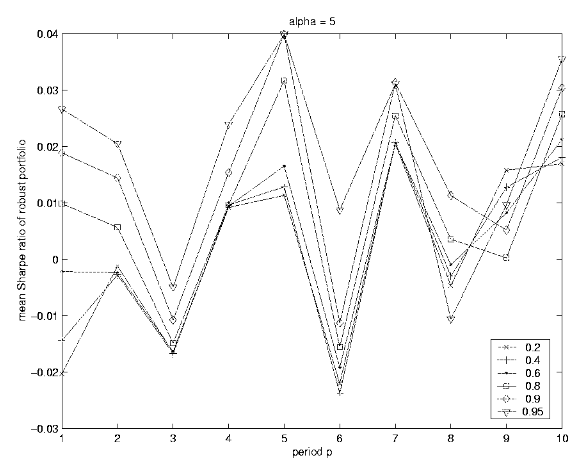




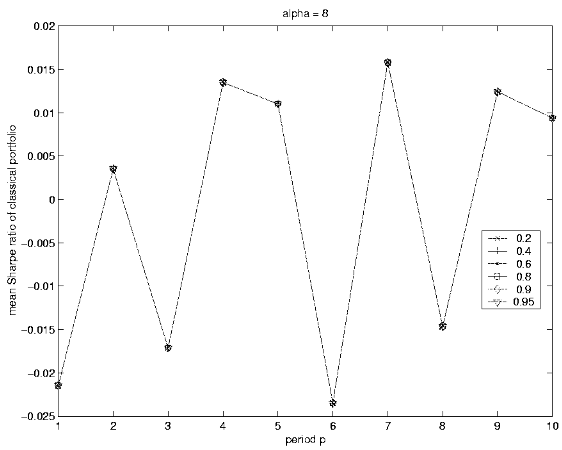


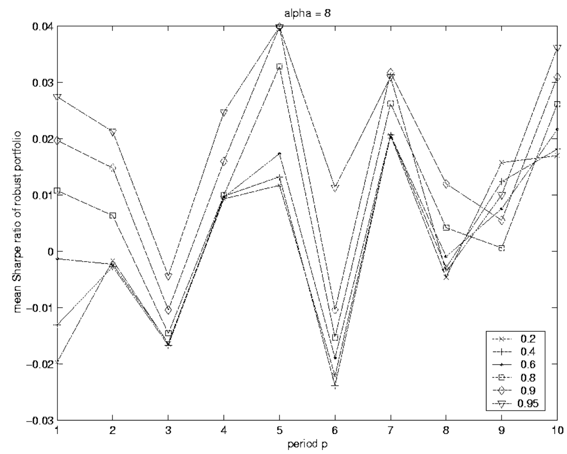


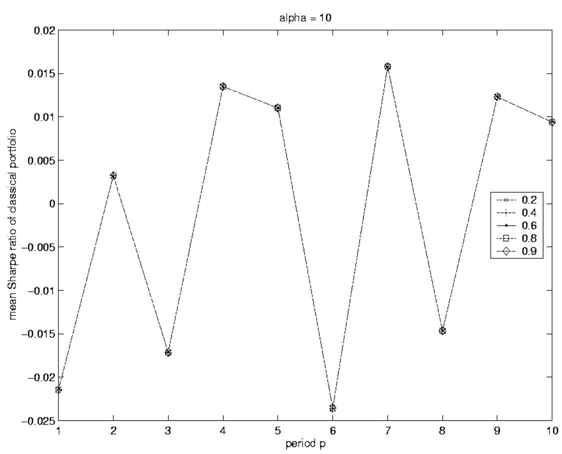


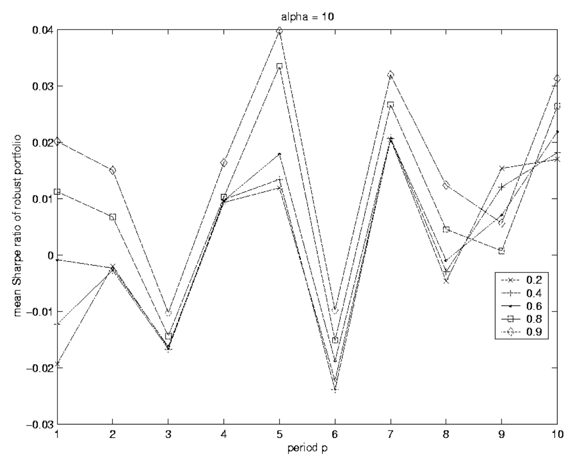


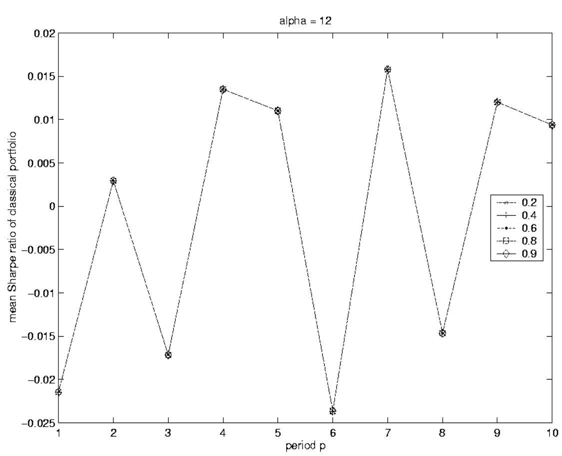
Data of Simulation 1
Mean Sharpe Ratio

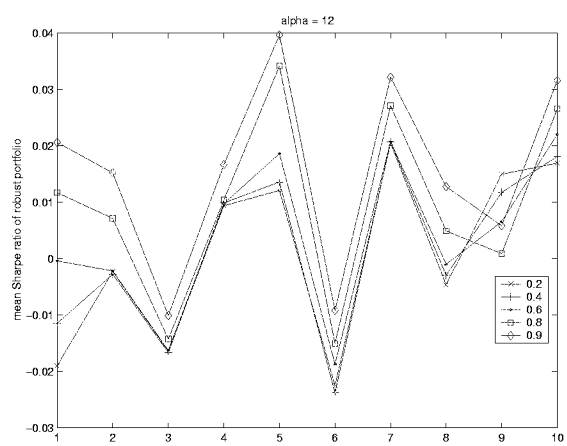




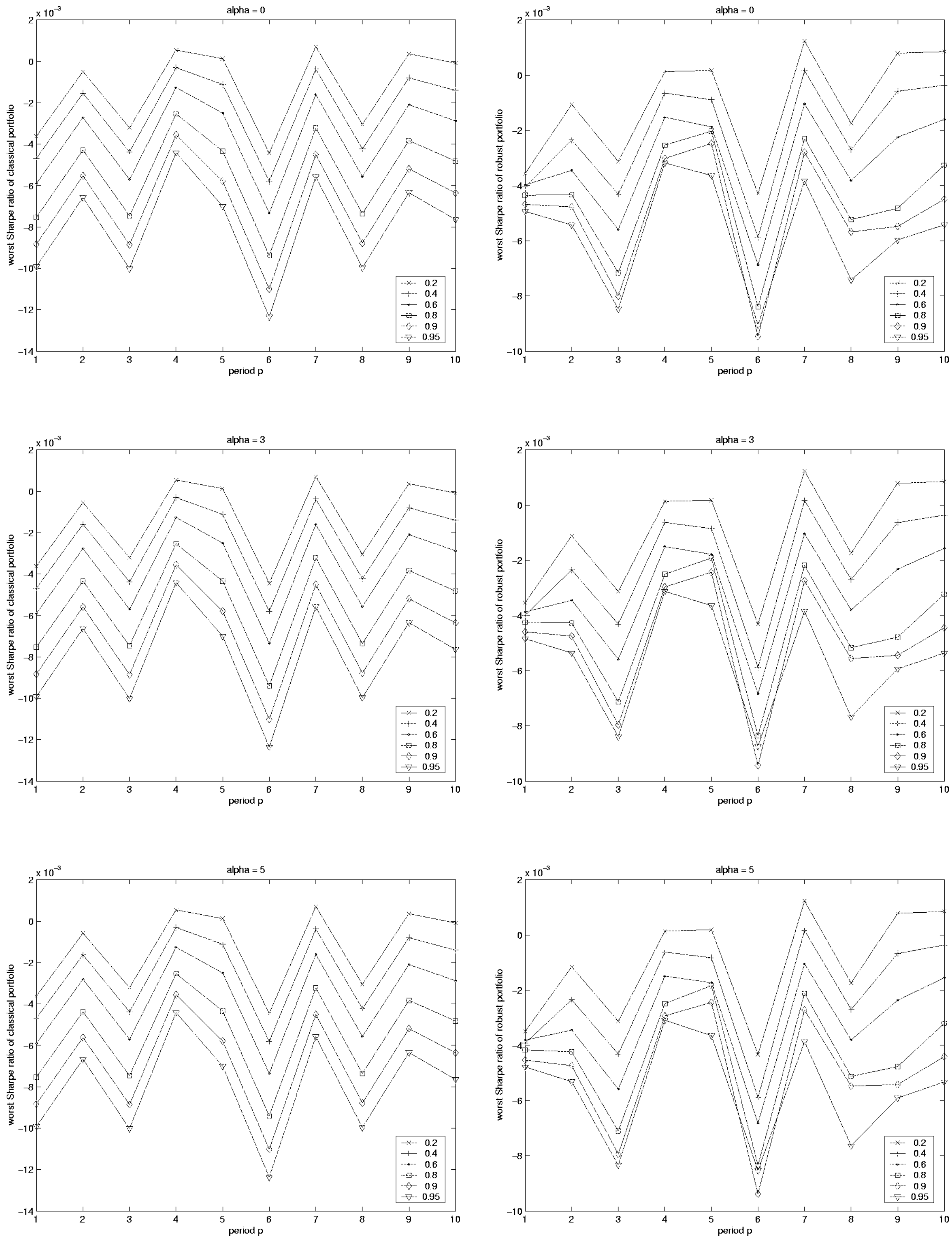




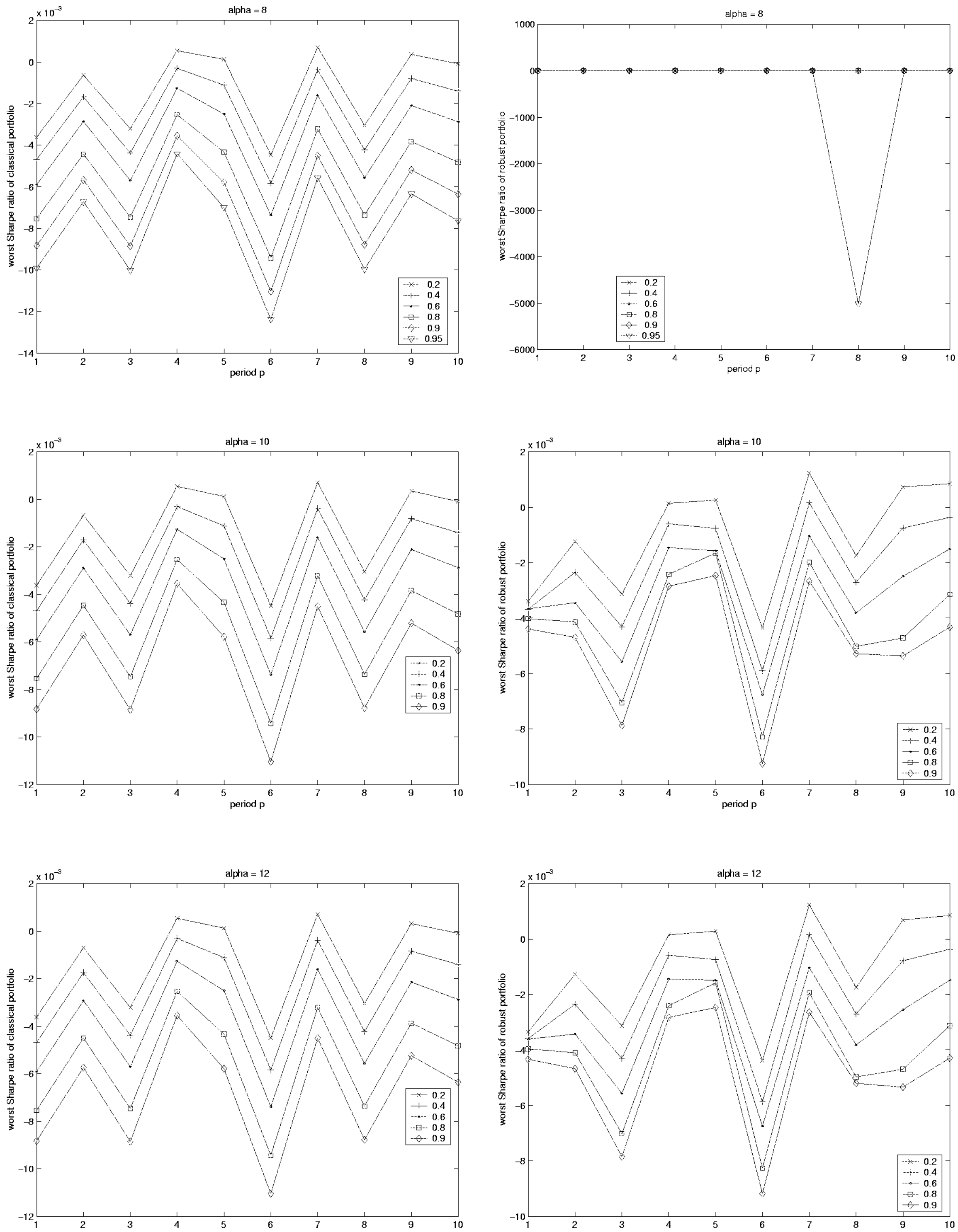




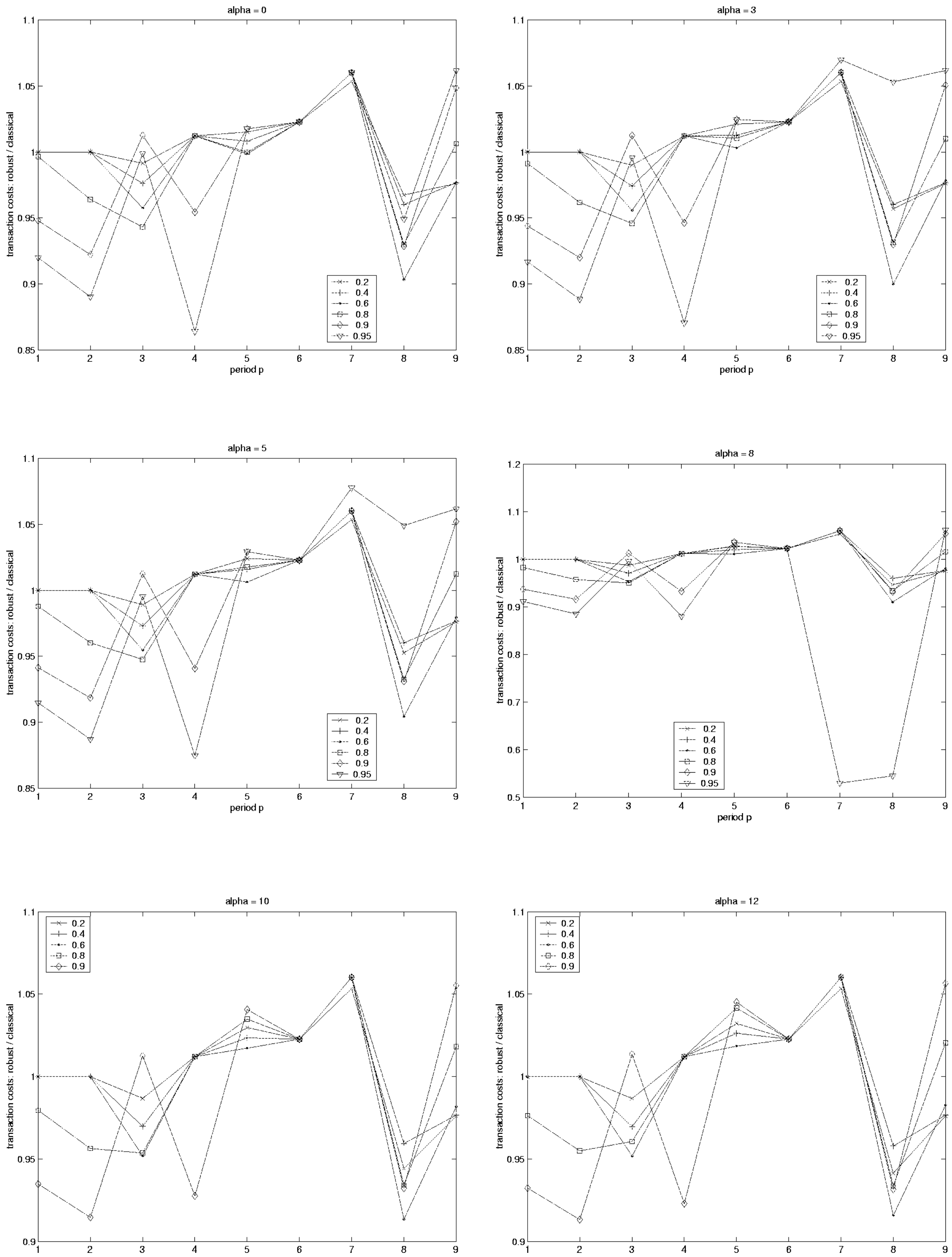
Data of Simulation 1
Worst Case Sharpe Ratio



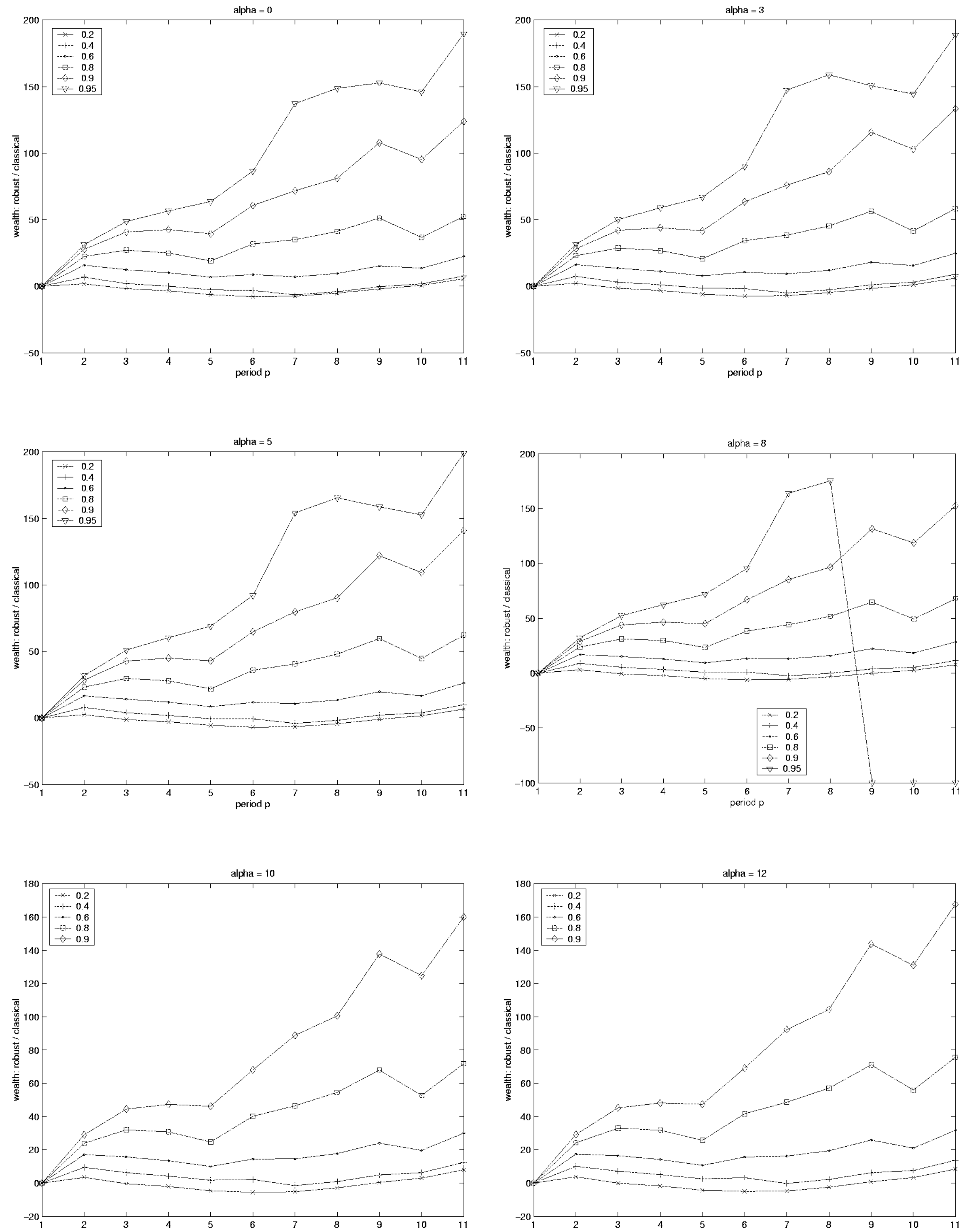
Data of Simulation 1
Worst Case Sharpe Ratio



DATA OF SIMULATION 1
RELATIVE TRANSACTION COSTS



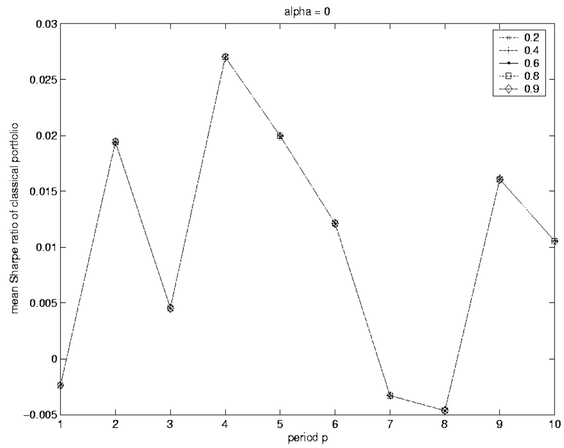
Data of Simulation 1
Mean Sharpe Ratio

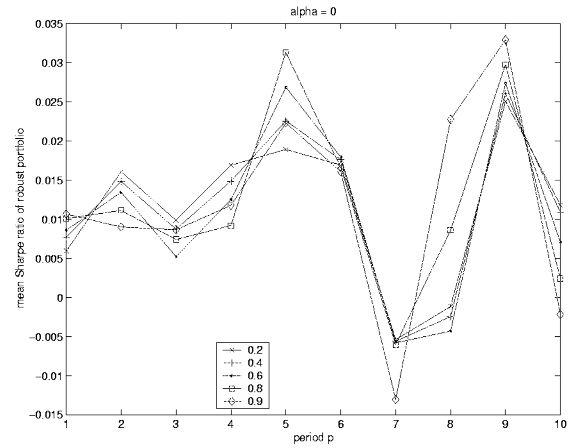


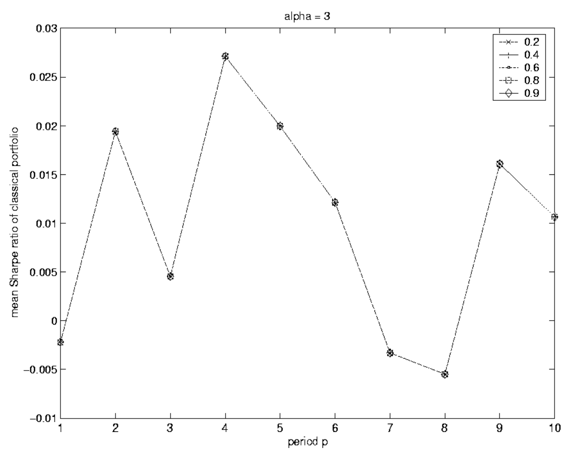
Appendix C

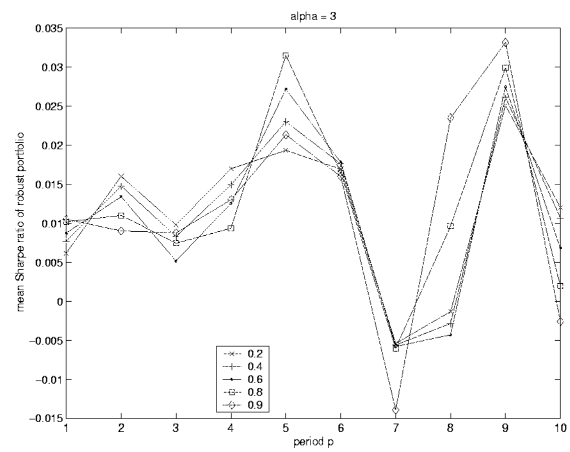
Graphical Results - Simulation 2

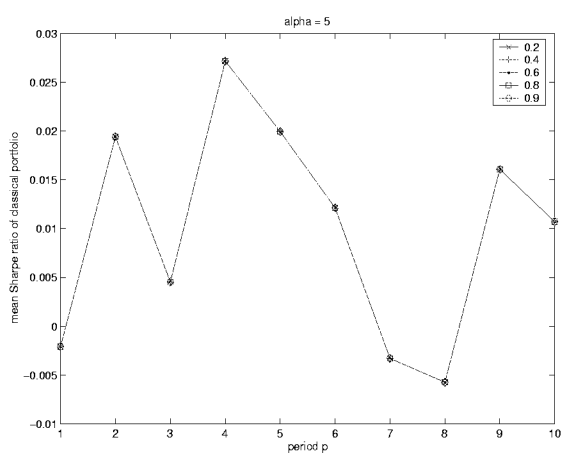
Data of Simulation 2
Mean Sharpe Ratio

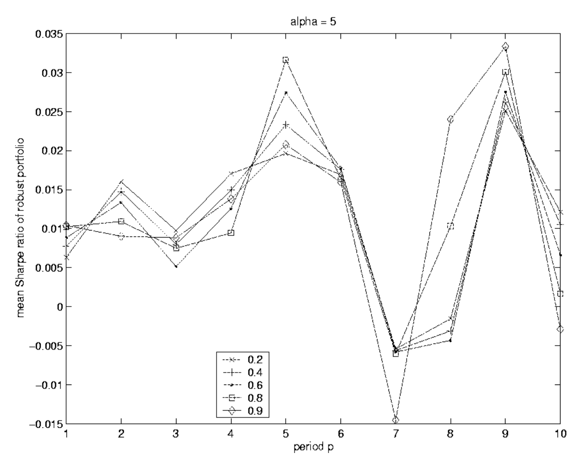




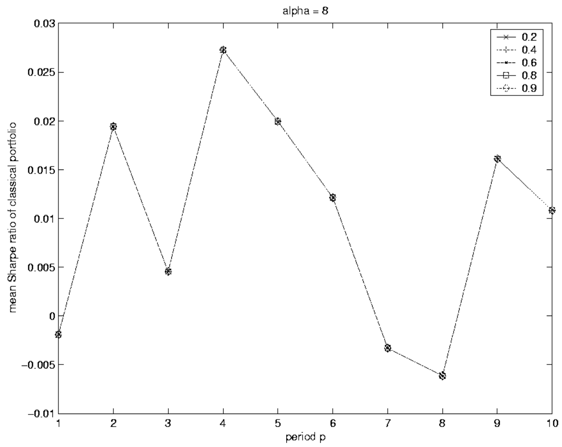


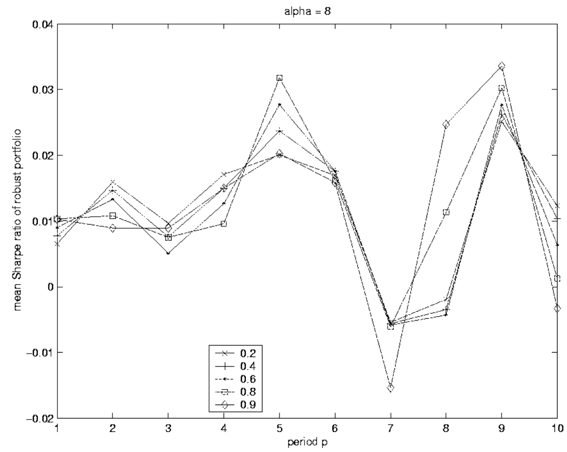


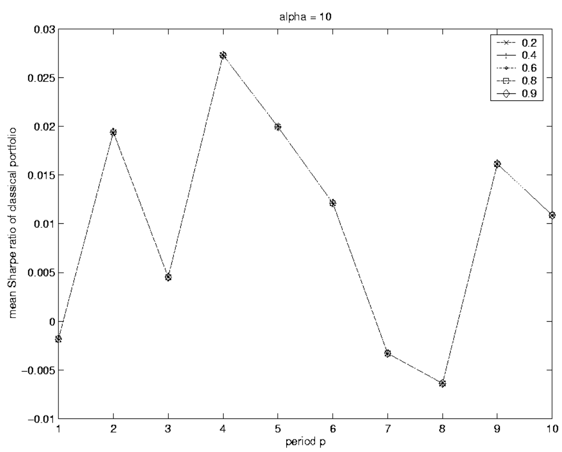


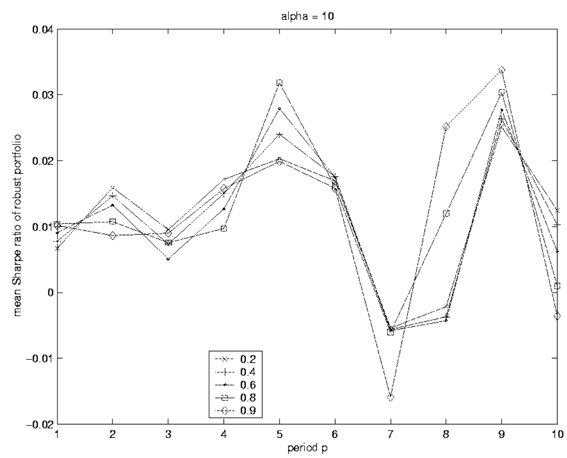


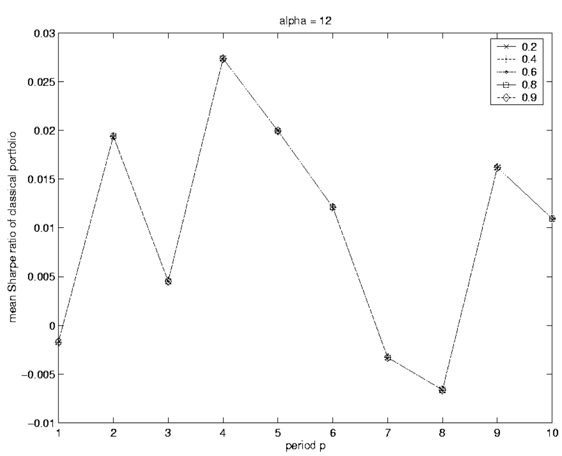
Data of Simulation 2
Mean Sharpe Ratio

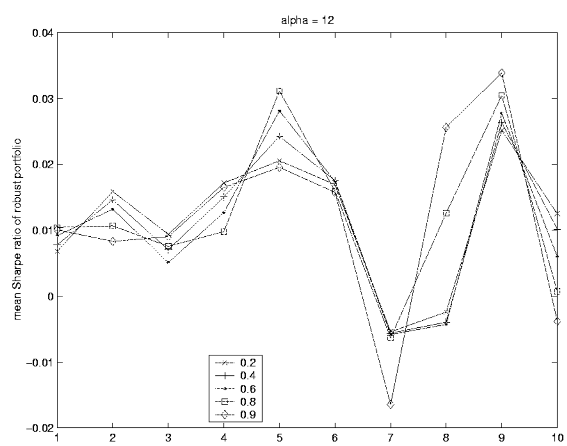




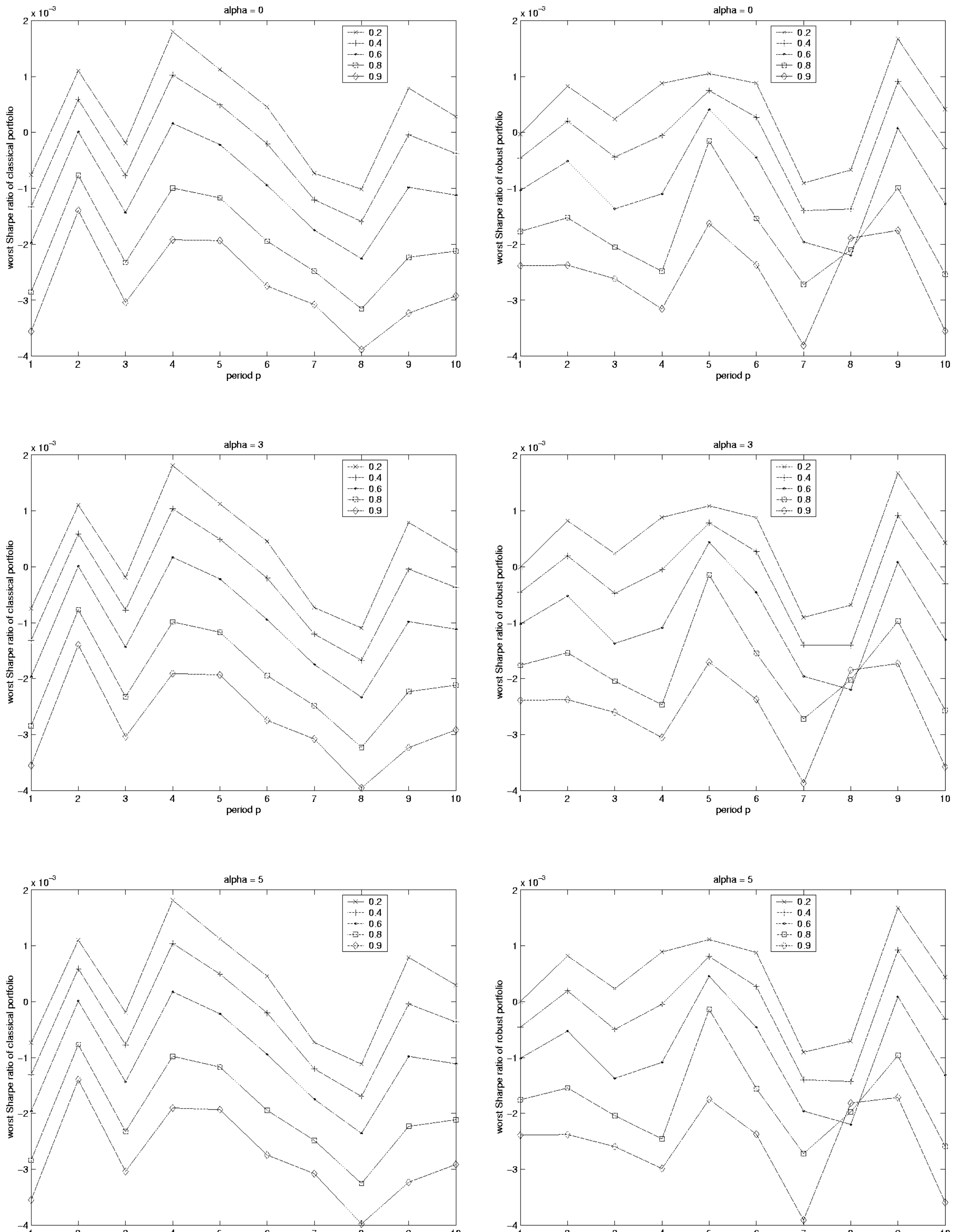




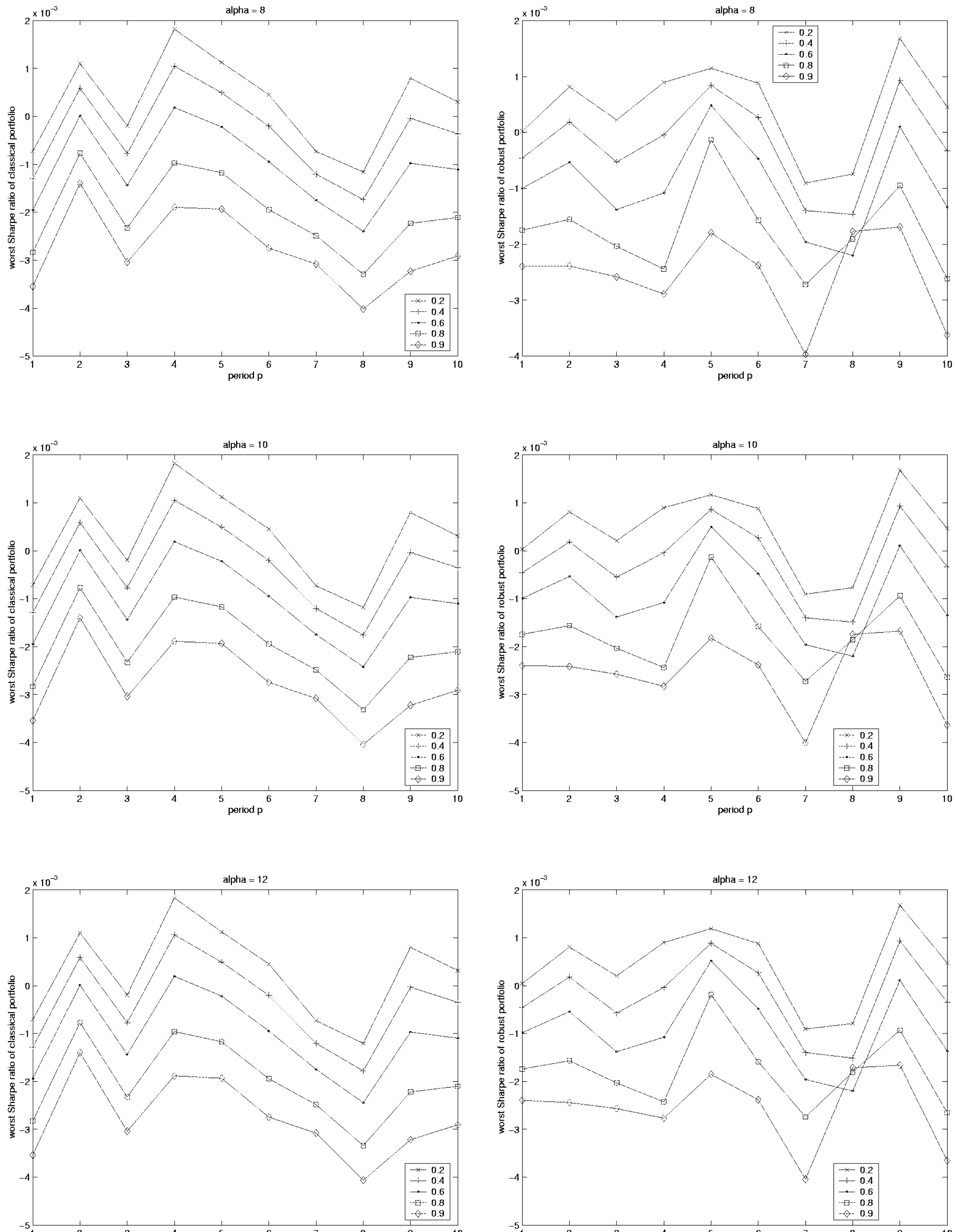




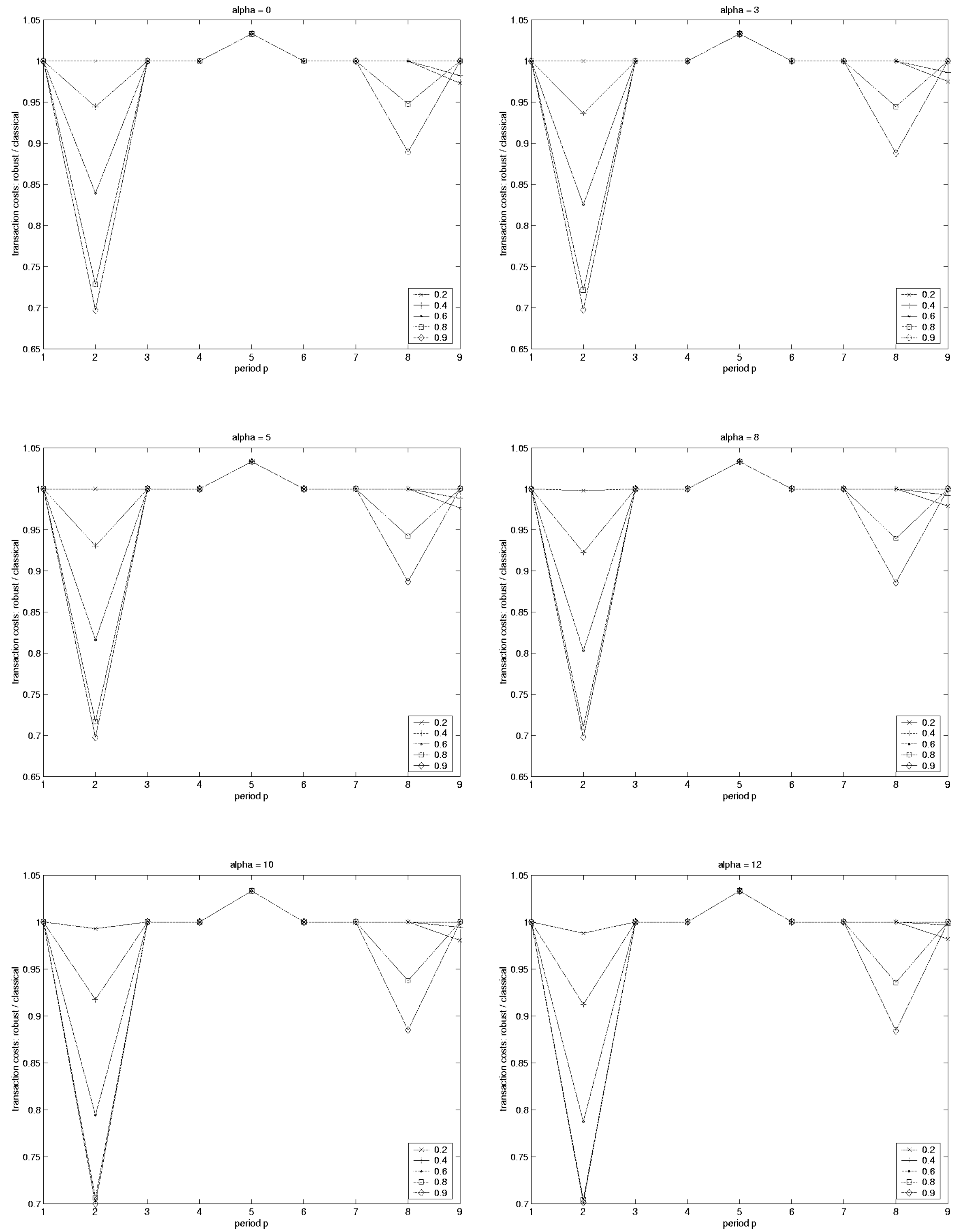
Data of Simulation 2
Worst Case Sharpe Ratio



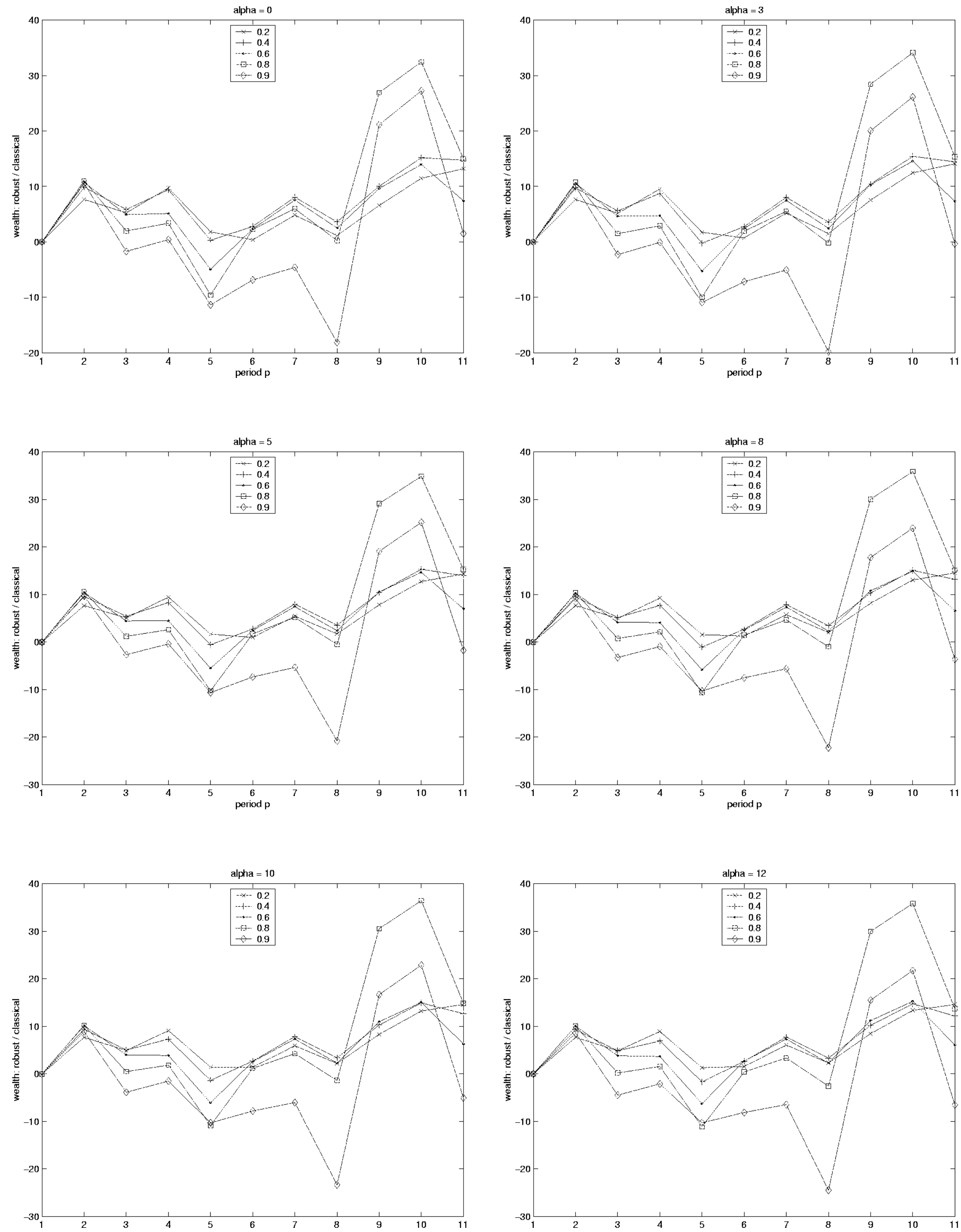
Data of Simulation 2
Worst Case Sharpe Ratio



Data of Simulation 2
Relative Transaction Costs



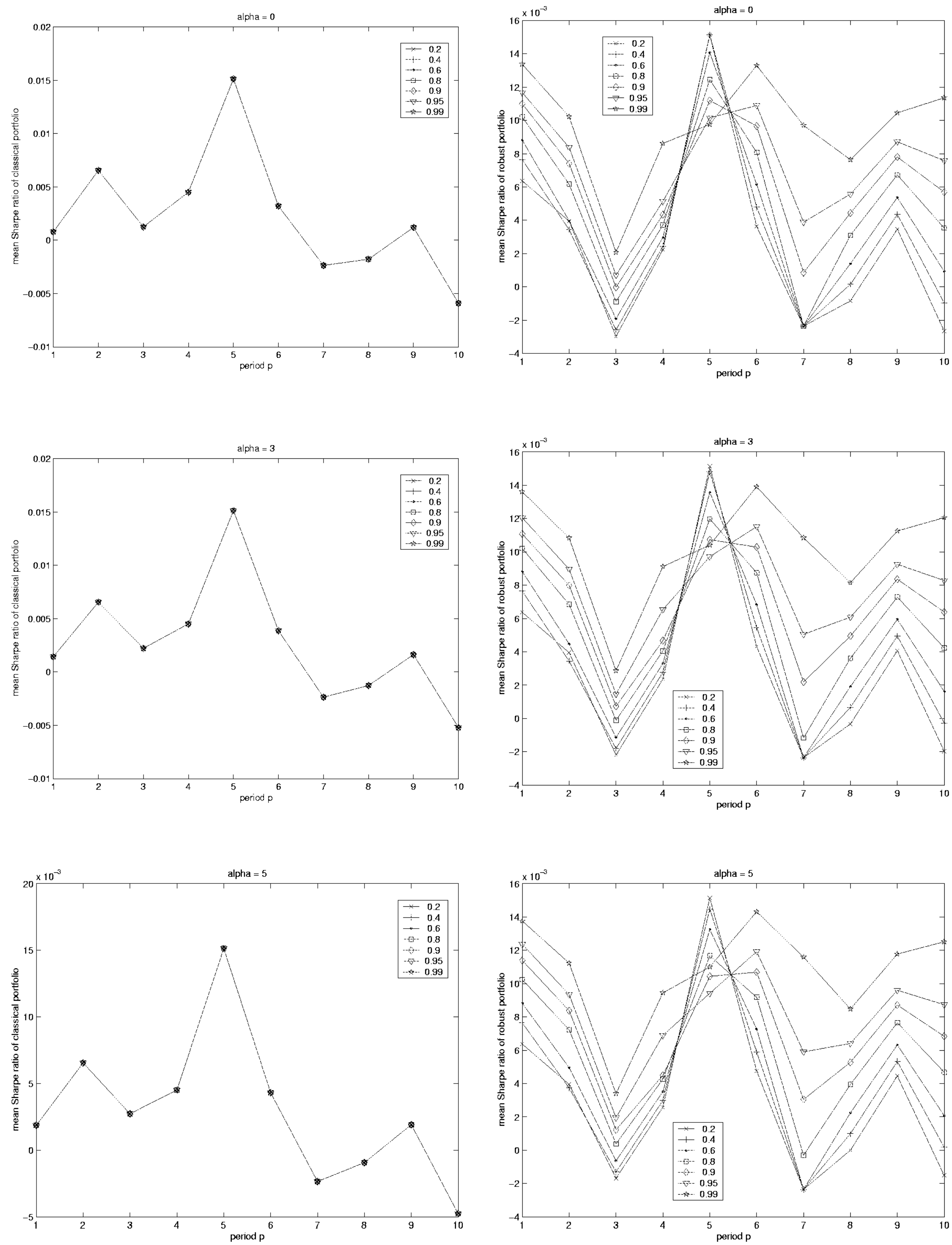
Data of Simulation 2
Relative Wealth



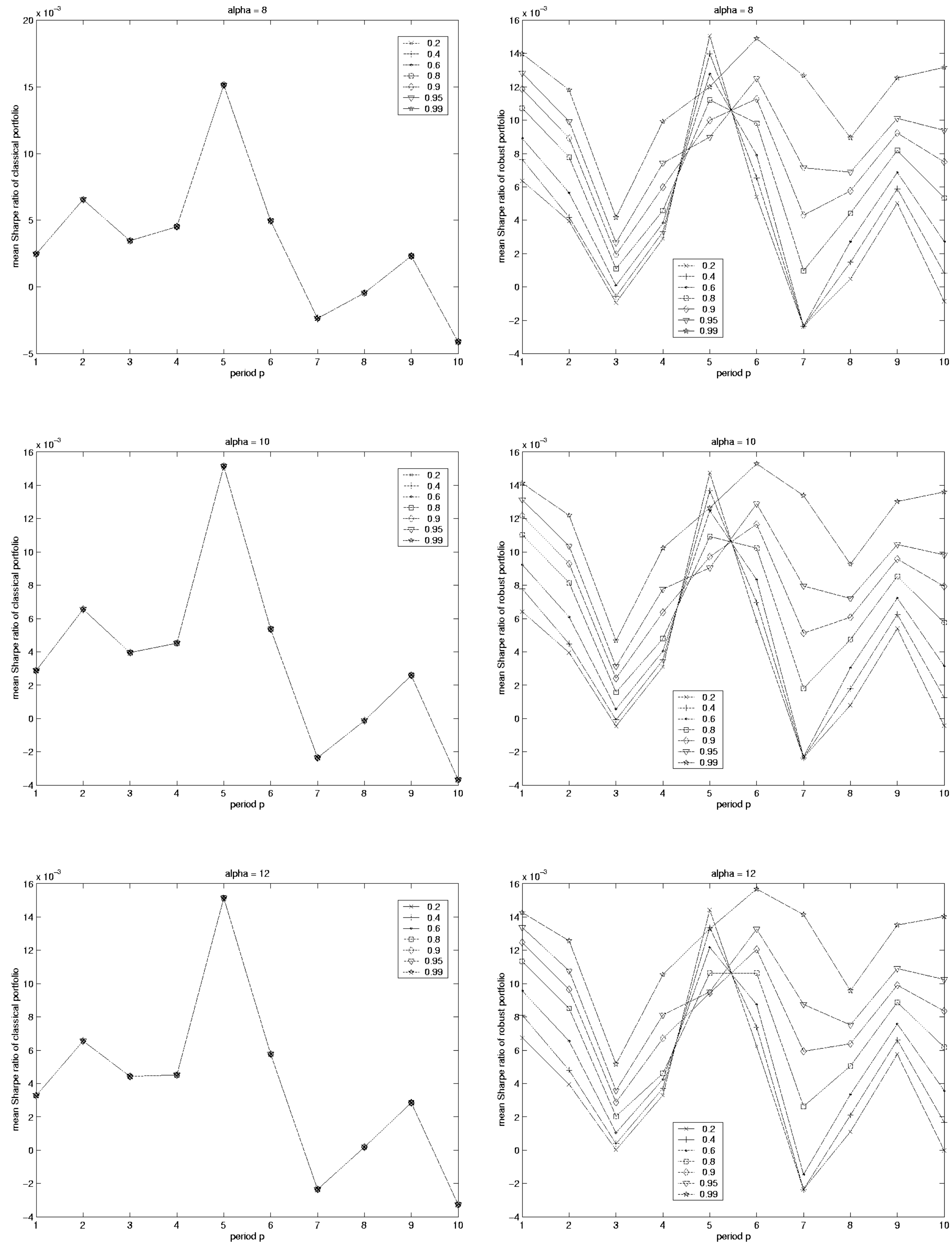
Appendix D

Graphical Results - Simulation 3

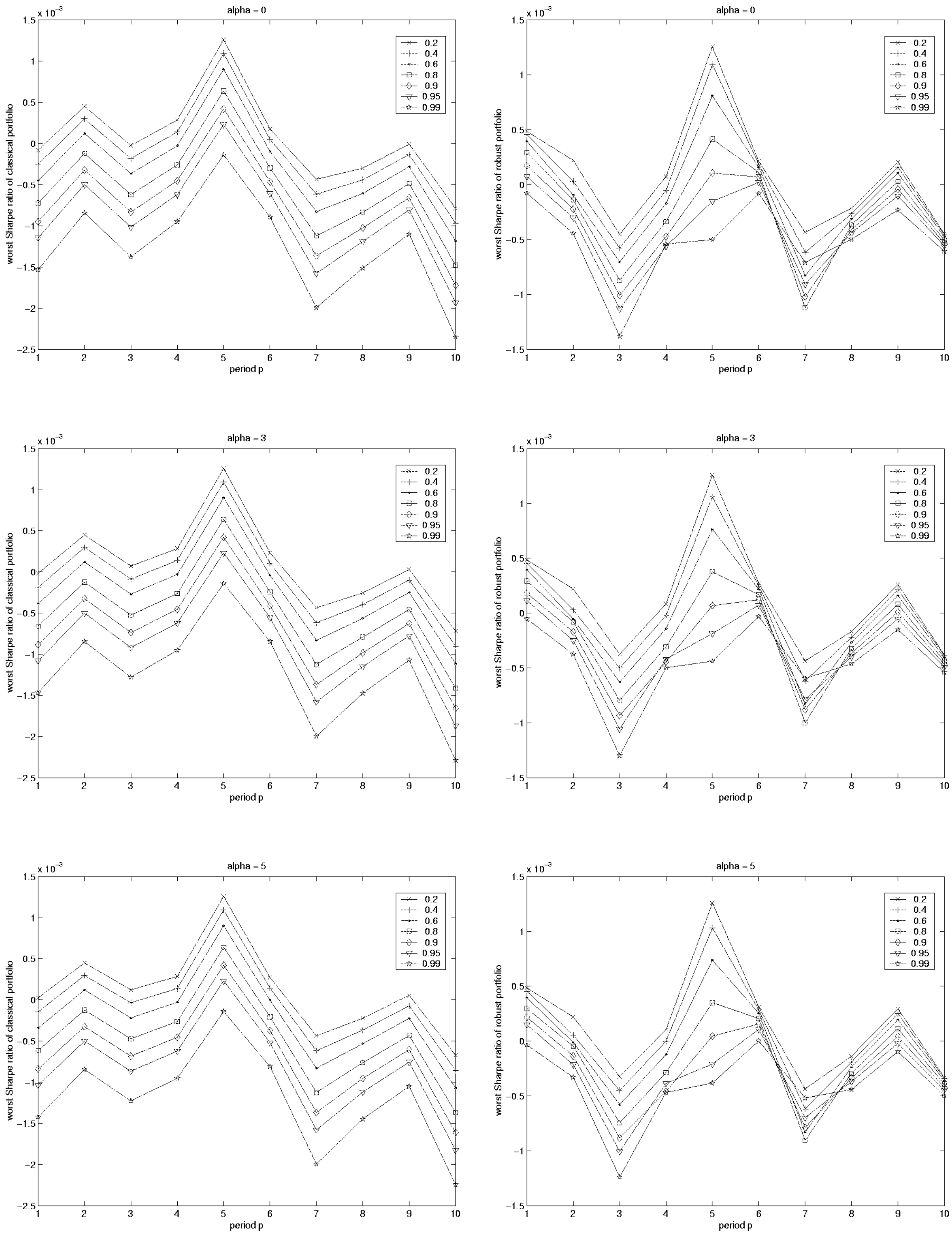
Data of Simulation 3
Mean Sharpe Ratio



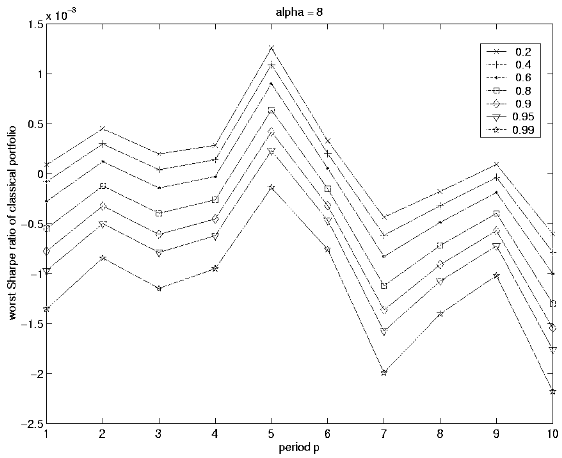
Data of Simulation 3
Mean Sharpe Ratio

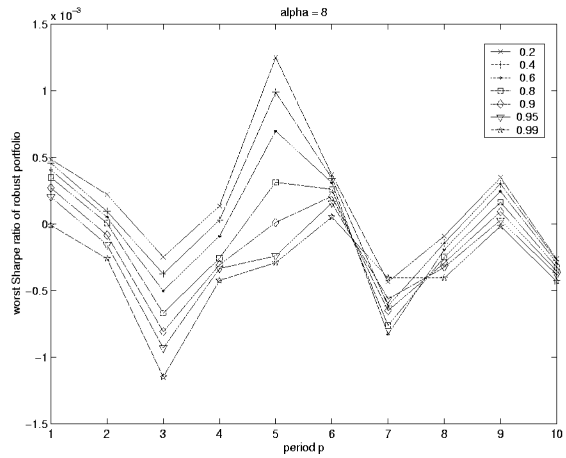


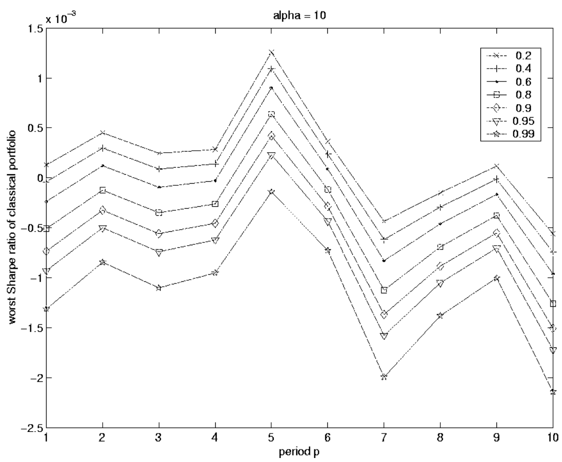
Data of Simulation 3
Worst Case Sharpe Ratio

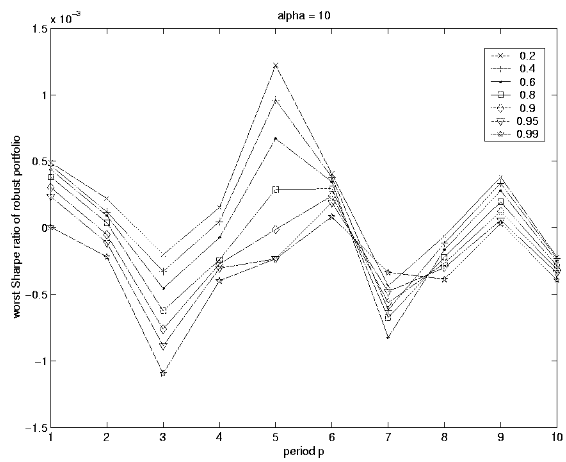


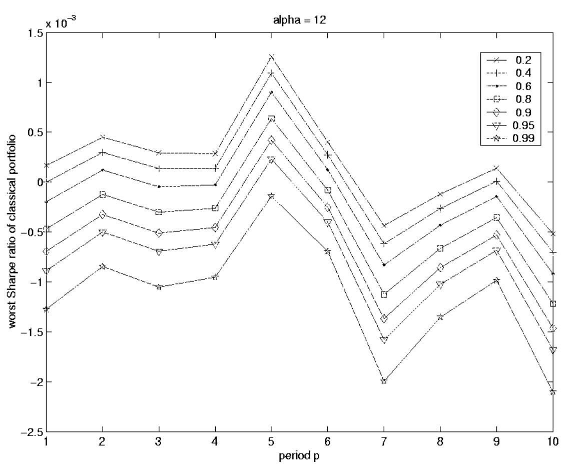
Data of Simulation 3
Worst Case Sharpe Ratio

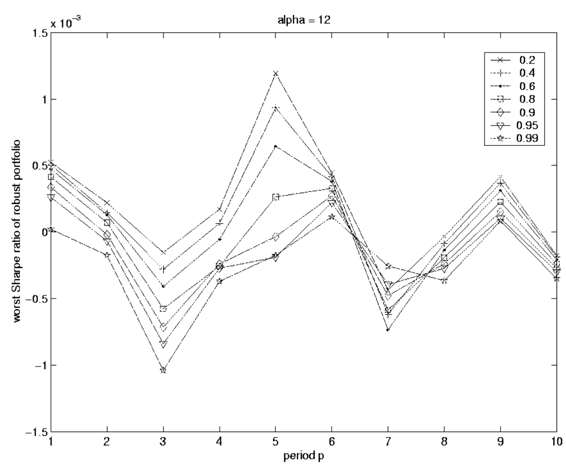




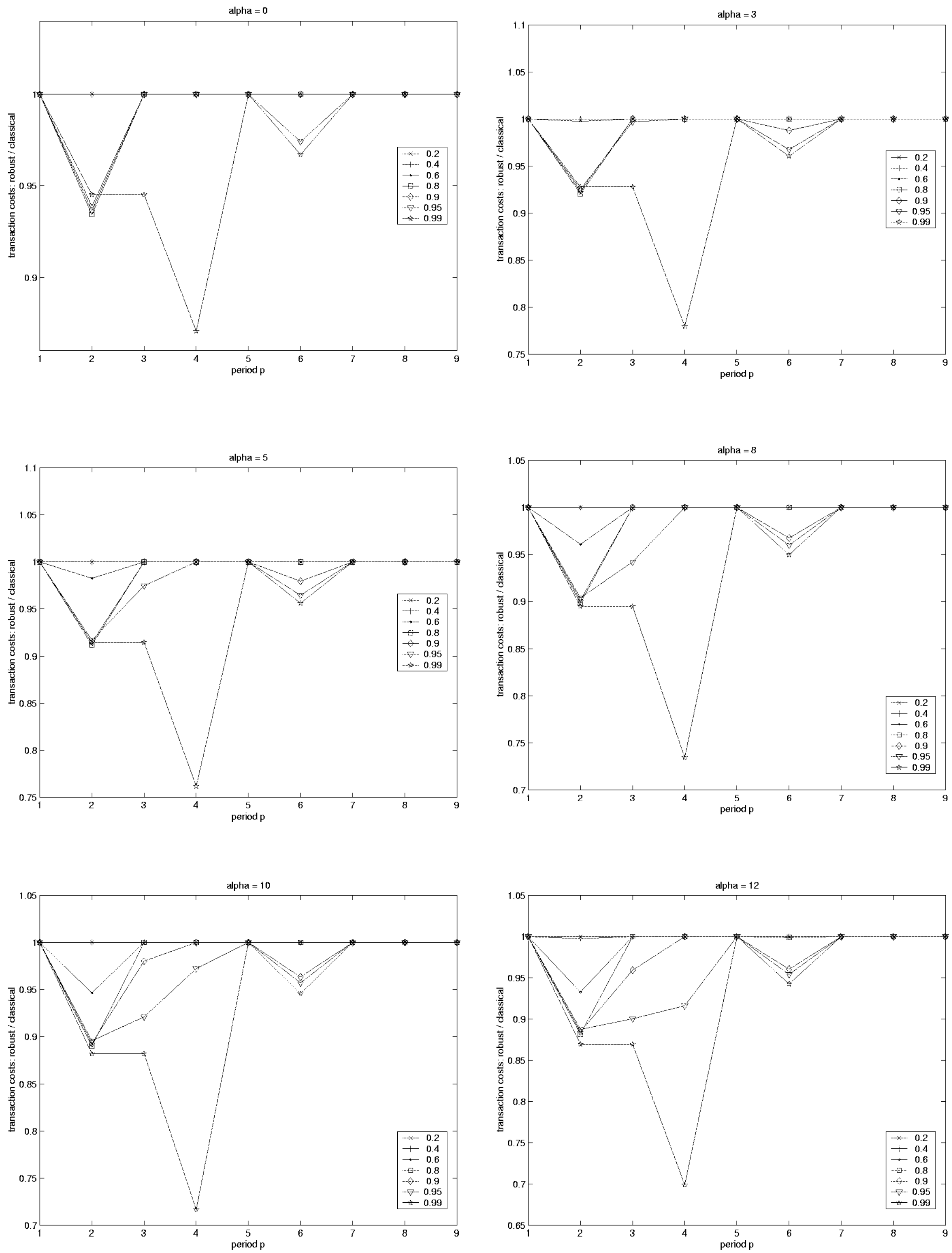




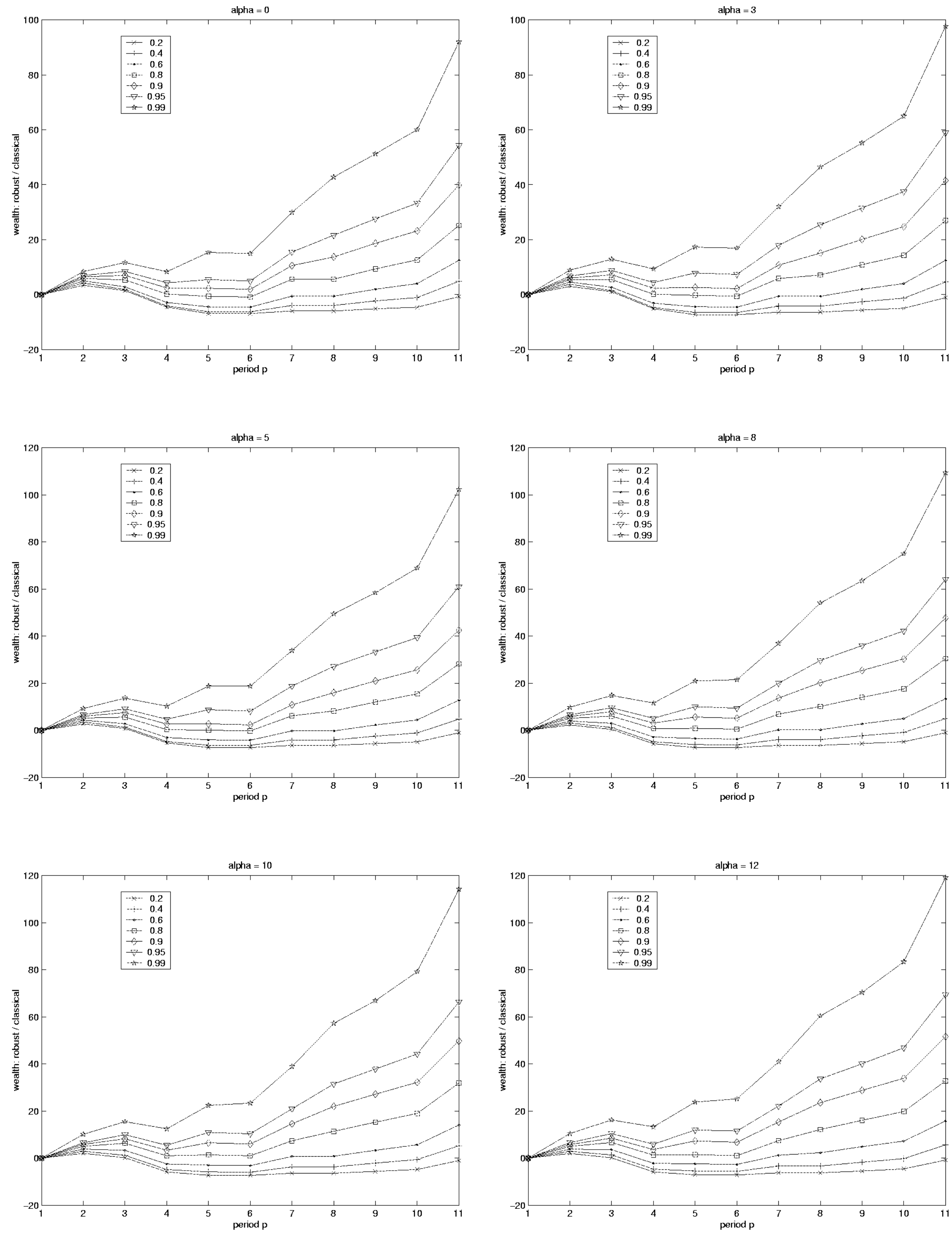




Data of Simulation 3
Relative Transaction Costs



Data of Simulation 3
Relative Wealth



Appendix E

MATLAB Programming Code

E.1 Calculation of ϕ_R and ϕ_C

```
% solves minimum variance problem with the classical and robust approach ---------
% calc (file,m )
% --- input parameters -----------------------------------------------------------
% file : name of file containing daily asset returns
% m    : number of eigenvectors used as factors

function calc(file,m);
info = fopen('INFO.dat','w'); % control file used to save Info of sedumi calls ---

% --------- read deaily asset returns from file ----------------------------------
assets = load(file);
assets = 100 * assets; % representation of the returns as percentage

% --------- initialize parameters ------------------------------------------------
k  = 90;   % sample size of each period
n  = size(assets,2);   % number of assets

% divide sample data into data sets with k samples and perform regression analyis
for p = 1:11

   Y = assets(k*(p-1) + 1 : k*p , : ); % Y : matrix of asset returns of period p
   cov_Y(:,:,p) = cov(Y);              % cov_Y: covariance matrix of asset returns

   % extract the eigenvectors of the m largest eigenvalues of cov_Y
```

```
   [v_Y,lambda_Y] = eigs(cov_Y(:,:,p),m);

   % calculate sample returns of these eigenvectors
   for l = 1:m
      v_Y(:,l) = v_Y(:,l)./sum(v_Y(:,l));
   end
   B(:,:,p) = [Y * v_Y]';       % B : matrix of factor returns (m x k)
   F(:,:,p) = cov(B(:,:,p)');   % F : estimated factor covariance matrix (m x m)

   % --- use regression analysis to calculate mu_0 and V_0 ----------------------
   A           = [ ones(k,1), B(:,:,p)'];
   X           = A\Y;
   mu_0(:,p)   = X(1,:)';
   V_0(:,:,p) = X(2:m+1,:);

   % --- calculate s_i^2 --------------------------------------------------------
   s2(:,p)   = [sum((Y-A*X).^2)/(k-m-1)]';

   % --- calculate inverse of A^T * A --------------------------------------------
   H_inv(:,:,p)   = inv(A'*A);
   cond_H(p) = cond(A'*A); % control variable

   % --- calculate cumulative returns for each asset of this period --------
   cum_Return(:,:,p)   = prod(Y./100 + ones(k,n));

   clear Y, v_Y, lambda_Y, l, A, X;
end

for alpha_i = [0,3,5,8,10,12,15]
   for omega_i = [20:20:80,90,95,99]

      % open file with name <alpha>_<omega>.dat to save results
      fid = fopen([int2str(alpha_i),'_',int2str(omega_i),'.dat'],'w');
      fprintf(fid,'\n alpha_i: %i \t omega_i: %i  \n',alpha_i,omega_i);
      fprintf(info,'\n alpha_i: %i \t omega_i: %i  \n',alpha_i,omega_i);
```

```
% recalculate alpha as daily return
alpha = (alpha_i/100 +1 )^(1/360) - 1;
alpha = alpha * 100;
omega = omega_i/100;

% --- critical values c_1 and c_m for the chosen level of omega  -----------
c_1 = finv(omega,1,k-m-1);
c_m = finv(omega,m,k-m-1);

for p = 1:11

   % ------ calculate gamma, rho and G for the robust calculation ----------
   gamma(:,p)  = sqrt(H_inv(1,1,p)*c_1*s2(:,p));
   G(:,:,p)    = B(:,:,p)*B(:,:,p)' - ...
                 1/k * [B(:,:,p) * ones(k,1)] * [B(:,:,p) * ones(k,1)]';
   rho(:,:,p)  = sqrt(m*c_m*s2(:,p));

   % --- calculate boundaries for d ---------------------------------------
   d(:,p)  = s2(:,p);

   % --- percentage of asset variance not explained by regression ----------
   check_d(p) = mean(d(:,p)./diag(V_0(:,:,p)' * F(:,:,p) * V_0(:,:,p)));

   % --- create input data for sedumi to calculate robust portfolio --------
   [AT_R,b_R,c_R,K_R] = mvp(alpha,mu_0(:,p),gamma(:,p),rho(:,p),d(:,p),...
                        F(:,:,p),G(:,:,p),V_0(:,:,p));

   % --- run sedumi to calculate robust portfolio --------------------------
   [X_R,Y_R,INFO_R(:,p)] = sedumi(AT_R,b_R,c_R,K_R);

   % --- create input data for sedumi to calculate classical portfolio -----
   [AT_C,b_C,c_C,K_C] = cmvp(alpha,mu_0(:,p),F(:,:,p),V_0(:,:,p),d(:,p));

   % --- run sedumi to calculate classical portfolio -----------------------
   [X_C,Y_C,INFO_C(:,p)] = sedumi(AT_C,b_C,c_C,K_C);
```

```
    % --- save calculated portfolio weights --------------------------------
    PHI_R(:,p)  = Y_R(3:83);
    PHI_C(:,p)  = Y_C(3:83);
    % --- save calculated portfolio variance --------------------------------
    Var_R(:,p) = Y_R(1:2);
    Var_C(:,p) = Y_C(1:2);
    % --- calculate expected return -----------------------------------------
    E_R(p)  = mu_0(:,p)' * PHI_R(:,p);
    E_C(p)  = mu_0(:,p)' * PHI_C(:,p);

end

% --- compare mean Sharpe Ratios ------------------------------------------
for p = 1:10
    mean_Sharpe_R(p) = mu_0(:,p+1)' * PHI_R(:,p) ...
        / sqrt(PHI_R(:,p)' * ...
        (V_0(:,:,p+1)' * F(:,:,p+1) * V_0(:,:,p+1) + diag(d(:,p+1))) ...
        * PHI_R(:,p)) ;
    mean_Sharpe_C(p) = mu_0(:,p+1)' * PHI_C(:,p) ...
        / sqrt(PHI_C(:,p)' * ...
        (V_0(:,:,p+1)' * F(:,:,p+1) * V_0(:,:,p+1) + diag(d(:,p+1))) * ...
        PHI_C(:,p)) ;
end
mean_Sharpe_rel = mean_Sharpe_R ./ mean_Sharpe_C

% --- calculate worst case Sharpe Ratio -----------------------------------
for p = 1:10

    % --- calculate worst case mu -------------------------------------------
    w_mu_R(p) = [mu_0(:,p+1) - gamma(:,p+1)]' * PHI_R(:,p);
    w_mu_C(p) = [mu_0(:,p+1) - gamma(:,p+1)]' * PHI_C(:,p);

    % --- calculate worst case variance by error term -----------------------
    w_d_R(p)  = PHI_R(:,p)'* diag(d(:,p+1)) * PHI_R(:,p);
    w_d_C(p)  = PHI_C(:,p)'* diag(d(:,p+1)) * PHI_C(:,p);
```

```
   % --- calculate worst case variance of the model -----------------------

   % --- create sedumi input and call sedumi ------------------------------
   % robust portfolio
   [AT_WR,b_WR,c_WR,K_WR] = wcsr(rho(:,p+1),F(:,:,p+1),G(:,:,p+1), ...
                                 V_0(:,:,p+1), PHI_R(:,p));
   [X_WR,Y_WR,INFO_WR(:,p)] = sedumi(AT_WR,b_WR,c_WR,K_WR);
   w_V_R(p)=Y_WR(1);

   % classical portfolio
   [AT_WC,b_WC,c_WC,K_WC] = wcsr(rho(:,p+1),F(:,:,p+1),G(:,:,p+1), ...
                                 V_0(:,:,p+1), PHI_C(:,p));
   [X_WC,Y_WC,INFO_WC(:,p)] = sedumi(AT_WC,b_WC,c_WC,K_WC);
   w_V_C(p)=Y_WC(1);

   % --- calculate worst case Sharpe ratio for period p --------------------
   worst_Sharpe_R(p) = w_mu_R(p)/(w_V_R(p) + w_d_R(p));
   worst_Sharpe_C(p) = w_mu_C(p)/(w_V_C(p) + w_d_C(p));

end
worst_Sharpe_rel = worst_Sharpe_R./worst_Sharpe_C;

% calculate transaction costs
R_T = sum(abs(PHI_R(:,2:10)-PHI_R(:,1:9)));
C_T = sum(abs(PHI_C(:,2:10)-PHI_C(:,1:9)));
T_rel = R_T ./ C_T;

% calculate wealth
W_R(1) = 1;
W_C(1) = 1;
for p = 1:10
   W_R(p+1) = cum_Return(:,:,p+1)* PHI_R(:,p) * W_R(p);
   W_C(p+1) = cum_Return(:,:,p+1)* PHI_C(:,p) * W_C(p);
end
W_rel = 100 * (W_R ./ W_C - ones(1,11));
```

```
    % --- print out the main results -----------------------------------------
    fprintf(fid,'\n %-13s: \t','mean_Sharpe_R');
    ... (original code contains more output lines)

    % --- print out sedumi control information ---------------------------------
    fprintf(info,'\n %-13s: \t','check_d');
    ... (original code contains more output lines)

  end
end
fclose(info);
```

E.2 Creation of dual standard form for SeDuMI

E.2.1 Classical Portfolio

```
% creates dual standard form for SeDuMi of classical minimum variance problem ----
% ------- decision vector - nu,delta,phi
% ------- dimensions        1  x 1  x n
function [At,b,c,K] = cmvp(alpha,mu_0,F,V_0,d);

% ------------- read in dimensions ---------------------------------------------
[m,n] = size(V_0);
% ------------- objective function ---------------------------------------------
b   = -sparse([1;1;zeros(n,1)]);

% ------------- equality -------------------------------------------------------
% ------------- 1^T * phi = 1 --------------------------------------------------
At  = sparse([0,0,ones(1,n)]);
c   = [1];
K.f = [1];

% ------------- nonnegativity constraints --------------------------------------
% ------------- nu, delta, phi >= 0 --------------------------------------------
At  = [At; ...
       -eye(2+n)];
c   = [c; zeros(2+n,1)];
```

```
K.l = 2+n;
% -------------- - alpha - [-mu_0^T] phi >= 0 -----------------------------------
At  = [At; ...
       0,0,[- mu_0']];
c   = [c; -alpha];
K.l = K.l + 1;

% -------------- Qcones -------------------------------------------------------
% -------------- (1 + delta,[2 D^1/2 phi; 1 - delta]) in Qcone -------------------
At  = [At; ...
      0,-1, zeros(1,n); ...
      zeros(n,2), -2*(diag(sqrt(d))); ...
      0,1, zeros(1,n)];
c   = [c; 1;zeros(n,1);1];
K.q = [n+2];
% -------------- (1 + delta,[2 D^1/2 phi; 1 - delta]) in Qcone -------------------
At  = [At; ...
       -1,0, zeros(1,n); ...
       zeros(n,2), -2*(V_0'*F*V_0)^0.5; ...
       1,0, zeros(1,n)];
c   = [c; 1;zeros(n,1);1];
K.q = [K.q,n+2];
```

E.2.2 Robust Portfolio

```
% creates dual standard form for SeDuMi of robust minimum variance problem -------
% ------- decision vector - nu,delta,phi,tau,sigma,t
% ------- dimensions        1  x 1  x n x 1 x  1  x m
function [At,b,c,K] = mvp(alpha,mu_0,gamma,rho,d,F,G,V_0);

% -------------- calculate matrices H, Q, Delta ----------------------------------
G_inv_sqrt = inv(G)^0.5;
H          = G_inv_sqrt * F * G_inv_sqrt;
[Q,Delta]  = schur(H);
% -------------- calculate eigenvalues of H --------------------------------------
lambda = eig(H);
L      = Q' * H^0.5 * G^0.5 * V_0;
```

```
% -------------- read in dimensions ----------------------------------------------
[m,n] = size(V_0);
% -------------- objective function ----------------------------------------------
b   = -sparse([1;1;zeros(n+2+m,1)]);

% -------------- equality --------------------------------------------------------
% -------------- 1^T * phi = 1 --------------------------------------------------
At  = sparse([0,0,ones(1,n),zeros(1,2+m)]);
c   = [1];
K.f = [1];

% -------------- nonnegativity constraints ---------------------------------------
% -------------- nu, delta, phi, tau, sigma, t >= 0 ------------------------------
At  = [At; ...
       -eye(4+n+m)];
c   = [c; zeros(4+n+m,1)];
K.l = 4+n+m;
% -------------- - alpha - [-mu_0^T + gamma^T] phi >= 0 --------------------------
At  = [At; ...
       0,0,[- mu_0' + gamma'],zeros(1,2+m)];
c   = [c; -alpha];
K.l = K.l + 1;
% -------------- - [- nu + tau + 1^T t] >= 0 -------------------------------------
At  = [At; ...
       -1, zeros(1,1+n),1,0,ones(1,m)];
c   = [c; 0];
K.l = K.l + 1;
%--------------- 1 - sigma lambda_max(H) >= 0 ------------------------------------
At  = [At; ...
       zeros(1,2+n+1),max(lambda),zeros(1,m)];
c   = [c; 1];
K.l = K.l + 1;

% -------------- Qcones ----------------------------------------------------------
% -------------- (1 + delta,[2 D^1/2 phi; 1 - delta]) in Qcone -------------------
At  = [At; ...
```

```
       0,-1, zeros(1,n+2+m); ...
       zeros(n,2), -2*diag(sqrt(d)), zeros(n,2+m); ...
       0,1, zeros(1,n+2+m)];
c   = [c; 1;zeros(n,1);1];
K.q = [n+2];
% ------------- ( sigma + tau, [ 2 rho^T \phi; sigma - tau ] ) in Qcone ---------
At  = [At; ...
       zeros(1,2+n),-1,-1, zeros(1,m); ...
       0,0,[-2 * rho'], zeros(1,2+m); ...
       zeros(1,2+n),1,-1, zeros(1,m)];
c   = [c;zeros(3,1)];
K.q = [K.q,3];
%-------------- ( 1-sigma*lambda_i+t_i, [ 2*w_i; 1-sigma*lamda_i-t_i ] in Qcone -
for i  = 1:m,
   At  = [At; ...
          zeros(1,2+n+1),lambda(i),zeros(1,i-1),-1,zeros(1,m-i); ...
          0,0,- 2 * L(i,:),zeros(1,2+m);...
          zeros(1,2+n+1),lambda(i),zeros(1,i-1), 1,zeros(1,m-i)
         ];
   c   = [c; 1 ; 0 ; 1];
   K.q = [K.q,3];
end
%----------------------------------------------------------------------------
```

E.2.3 Worst case Sharpe ratio

```
% creates dual standard form for SeDuMi worst case Sharpe Ratio
% [At,b,c,K] = wcsr(rho,F,G,V_0,phi);
% ------- decision vector - nu,tau,sigma,t
% ------- dimensions        1  x 1 x  1  x m


function [At,b,c,K] = wcsr(rho,F,G,V_0,phi);


% ------------- calculate matrices H, Q, Delta ---------------------------------
G_inv_sqrt = inv(G)^0.5;
H          = G_inv_sqrt * F * G_inv_sqrt;
```

```
[Q,Delta]  = schur(H);
% ------------- calculate eigenvalues of H ----------------------------------
lambda = eig(H);
L      = Q' * H^0.5 * G^0.5 * V_0;
% ------------- read in dimensions ---------------------------------------------
[m,n] = size(V_0);


% ------------- objective function ---------------------------------------------
b   = -sparse([1;zeros(2+m,1)]);

% ------------- nonnegativity constraints --------------------------------------
% ------------- nu, tau, sigma, t >= 0 ------------------------------
At  = [ -eye(3+m)];
c   = [ zeros(3+m,1)];
K.l = 3+m;
% ------------- - [- nu + tau + 1^T t] >= 0 -----------------------------------
At  = [At; ...
       -1,1,0,ones(1,m)];
c   = [c; 0];
K.l = K.l + 1;
%-------------- 1 - sigma lambda_max(H) >= 0 ----------------------------------
At  = [At; ...
       zeros(1,2),max(lambda),zeros(1,m)];
c   = [c; 1];
K.l = K.l + 1;

% ------------- Qcones ----------------------------------------------------------
% ------------- ( sigma + tau, [ 2 rho^T \phi; sigma - tau ] ) in Qcone ---------
At  = [At; ...
       zeros(1,1),-1,-1, zeros(1,m); ...
       0,zeros(1,2+m); ...
       zeros(1,1),1,-1, zeros(1,m)];
c   = [c;0;2*rho'*phi;0];
K.q = [3];
%-------------- ( 1-sigma*lambda_i+t_i, [ 2*w_i; 1-sigma*lamda_i-t_i ] in Qcone -
```

```
for i  = 1:m,
   At  = [At; ...
          zeros(1,2),lambda(i),zeros(1,i-1),-1,zeros(1,m-i); ...
          0,zeros(1,2+m);...
          zeros(1,2),lambda(i),zeros(1,i-1), 1,zeros(1,m-i)
         ];
   c   = [c; 1 ; 2 * L(i,:)* phi ; 1];
   K.q = [K.q,3];
end
%---------------------------------------------------------------------------------
```

References

[1] E. Elton and N. Gruber. *Modern Portfolio Theory and Investment Analysis.* John Wiley & Sons, 1995.

[2] D. Goldfarb and G. Iyengar. Robust portfolio selection problems. *CORC REPORT 2002-03*, December 2001. Revised: May 25th, 2002, July 24th, 2002.

[3] M. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret. Applications of second-order cone programming. *Linear Algebra and Applications*, 284(1:3):193–228, 1998.

[4] Falko Lorenz. *Lineare Algebra II.* B-I-Wissenschaftsverlag, 1992.

[5] H. Markowitz. Portfolio selection. *Journal of Finance*, 7:77, 1952.

[6] Pindyck Robert, S. and L. Rubinfeld Daniel. *Econometric Models and Econometric Forecasts.* McGRAW-HILL International Editions, 1998.

[7] F. Sturm, J. Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones (updated for version 1.05). *Optimization Methods and Software*, 11-12:625–653, 1999.