MEAN-RISK PORTFOLIO OPTIMIZATION PROBLEMS WITH RISK-ADJUSTED MEASURES

BY NAOMI LIORA MILLER

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Operations Research

Written under the direction of

Andrzej Ruszczyński

and approved by

New Brunswick, New Jersey

October, 2008

ABSTRACT OF THE DISSERTATION

Mean-Risk Portfolio optimization problems with risk-adjusted measures

by Naomi Liora Miller

Dissertation Director: Andrzej Ruszczyński

We consider the problem of optimizing a portfolio of finitely many assets whose returns are described by a joint discrete distribution. We formulate the mean-risk model, using as risk functions the semideviation and weighted deviation from quantile. Using representation theorems from convex analysis, we write the portfolio problem equivalently as a zero-sum matrix game, and provide convex optimization techniques for its solution. A new set of risk-adjusted probability measures is derived from the optimal saddle point solution of the game.

The risk-adjusted probability measures can be used to evaluate portfolio performance. An illustrative example is provided in which these measures are derived for a portfolio of 200 assets, and are used to evaluate a market portfolio and optimal risk-averse portfolio. The results suggest the mean-risk portfolio is more robust than a market portfolio.

We extend the above mean-risk model to the two-stage portfolio problem, where there are two investment periods and the option to rebalance inbetween. The resulting model is a two-stage stochastic programming problem, with mean-risk objectives in

each stage. First and second stage risk-adjusted probability measures are derived in a similar fashion to the one investment period problem.

Using as risk functions semideviation and weighted deviation from quantile in both stages, we calculate the risk adjusted measures in a numerical example with 100 assets. These measures are used to evaluate a two-stage market portfolio and optimal risk-averse portfolio.

We extend the cutting-plane and multi-cut algorithms for solving linear two-stage stochastic problems to the two-stage mean-risk portfolio problem. The two-stage portfolio problem is also formulated as one large linear program. We provide an illustrative example, where a two-stage portfolio problem with risk functions semideviation and weighted deviation from quantile is solved, using these two methods and the simplex method. The performance of these three methods is compared for solving the portfolio problem. On large examples, the extended cutting-plane and multi-cut plane algorithms solve where the linear program fails.

Acknowledgements

I came to Rutgers University in the fall of 2004 to pursue a Ph.D in Operations Research. This was both an exciting and challenging time. Exciting as I was to expand my knowledge and live away from home for the first time. Challenging as I was in a strange place with almost no friends or family. The inevitable homesickness would creep in. One of the few people I did know at the time, Al and Carol, would keep company on the weekends and take me shopping when I didn't have a car. Their friendship in the first year was very helpful. I also acknowledge Ellen and Jerry Cohen, and Noam and Michal, who invited me to Jewish seders.

I moved to a house on North Second Ave. in the late summer of 2005, where I met my new roommates, Anupama Reddy and Rohan Fernandes. This was my first real exposure to Indian cooking and movies. Their guidance through graduate school was invaluable. Other people to mention in this regard include Lilya Fedzhora, Mine Subasi and Vimla Gulabani.

It was in this second year of graduate school that I also began an indepedent research study with my future advisor, Andrzej Ruszczynski. We began with the topic of risk-adjusted bond valuation and ended up with the current topic of portfolio optimization and risk-adjusted probability measures. I have enjoyed working with him these last three years, and his guidance was invaluable.

I want to thank Dr. Eckstein for his helpful corrections and suggestions for the dissertation draft.

Other friends in which I have spent much time and enjoyment with include Aritanan Gruber, Minh Pham, Natalia Santa maria Tobar, Bjita Majumdar, Lilya Fedzhora and Shilpa Shanbag.

The support of my family and friends from Toronto was invaluable. I mention here my parents, Harold Roy Miller, Jo-anne Miller, brother Jonathan Miller and Aunt Shirley Miller. Also my very close friends Natalya Neman, Lee-Anne Wachtel and Lindsay Shorser.

I make special mention of Rohan Fernandes, for his companionship and guidance throughout graduate school.

Dedication

To my late grandparents, Isidore Jack Miller and Lily Miller

Table of Contents

List of Tables

List of Figures

Chapter 1

Preliminaries

1.1 Introduction

The problem of optimizing a portfolio of finitely many assets is a central problem in theoretical finance. Markowitz introduced the classical approach to this problem in his seminal paper [18]. In it, he argued that the portfolio performance should be measured in two distinct dimensions: the mean $E[R]$ of the portfolio return R, and the risk $r[R]$, which measures the variation of the return. In the mean-risk approach, the objective was to select from the universe of all portfolios those that are efficient: for a fixed value of the mean, the risk is minimized, and for a fixed value of the risk, the mean is maximized.

The mean-risk approach has many advantages: it allows a trade off analysis between mean and risk. Moreover, it allows one to formulate the portfolio problem as a parametric optimization problem.

The question of what risk function to use in the mean-risk approach has been examined extensively in the literature. One important direction of research was initiated by Artzner *et al* [2] in their paper "Coherent Risk Measures". In it, they outlined a set of mathematical properties that a meaningful risk measure ought to satisfy. It was argued that these axioms reflect the interests of risk-averse investors. In another vein, Ogryczak and Ruszczyński [22, 23, 24] used stochastic dominance relations to compare portfolio returns. They identified several risk functions for which the optimal portfolio returns are non-dominated in terms of the second order stochastic dominance relation. Important examples include the semideviation and weighted deviations from quantile.

Another important area of research is the optimization of a portfolio over multiple investment periods. In particular, what optimization models should be used, and when or if to rebalance a portfolio. An important recent development in this area has been the conditional risk mapping approach [37]. The idea is to develop a model in which information from the previous investment period can be used in the decision for the next investment period. In the conditional risk mapping approach, such information is incorporated using a stochastic programming formulation.

In the first part of this dissertation, we examine the mean-risk portfolio problem for coherent risk functionals. We begin with a formal description of the portfolio problem, followed by a literature review. In chapter 2, we review the main representation theorem for coherent risk measures and show that several mean-risk objective functions are coherent. This, in combination with the optimality and duality theorems for the portfolio problem, allow the mean-risk portfolio problem to be written as a zero-sum matrix game. In [21], it is proved that certain probability measures arise as part of the optimal saddle point solution to the game. We call these measures *risk-adjusted probability measures*. In chapter 2, convex optimization techniques are provided for solving the mean-risk portfolio problem with coherent risk functions in the form of semideviation and weighted deviation from quantile. Closed forms for the risk-adjusted probability measures are constructed for the above mean-risk models in these sections.

In chapter 3, the mean-risk portfolio problems for risk measures mean-semideviation and mean-weighted deviation from quantile are solved for a portfolio of 200 assets. The risk-adjusted probability measures are constructed for these examples. A market portfolio is also constructed in each case and compared to the optimal mean-risk portfolio.

In the second part of the dissertation, we examine two-stage portfolio problems. We begin with a formal description of the two-stage portfolio problem in chapter 4. We review the conditional risk mapping approach to two stage optimization problems and develop the two-stage mean risk model from it. In chapter 4, we argue for the use of the conditional risk mapping approach and introduce the property of time consistency, which this method satisfies. In chapter 5, we extend optimality and duality theory from the first section to composite coherent risk measures, and develop matrix game representations for both first and second stage optimization problems. We derive first and second stage risk adjusted probability measures as part of the optimal saddle point solutions to these problems. For the risk functions defined as semideviation and weighted deviation from quantile, we derive closed forms for the first and second stage risk-adjusted probability measures.

In chapter 6, we review Benders' decomposition method for solving a linear two stage stochastic programming problem. This method is extended to the two-stage mean risk model, in particular, for the risk measures semideviation and weighted deviation from quantile. The multi-cut version of Benders' decomposition is introduced and extended in a similar manner. A large linear programming formulation of the two stage mean-risk portfolio problems for the risk measures semideviation and weighted deviation from quantile is given.

Using semideviation and weighted deviation from quantile as risk functions in both stages, we calculate the risk adjusted measures in a numerical example with 100 assets. These measures are used to evaluate a two-stage market portfolio and optimal risk-averse portfolio. In chapter 7, we solve a two-stage portfolio problem with risk functions semideviation and weighted deviation from quantile, using these two methods and the simplex method. The performance of these three methods is compared for solving the portfolio problem. On large examples, the extended cutting-plane and multi-cut algorithms solve where the linear program fails.

1.2 The Portfolio Problem

We begin with a formal description of the portfolio problem. Consider a collection of n assets in which we would like to invest some intial capital C. For simplicity, we will take $C = 1$. The n-dimensional vector R represents the collection of asset returns, with each component R^j equal to the return of asset j, for $j = 1..n$. R is assumed to be an integrable random variable on given probability space $(\Omega, \mathcal{F}, P)$, with $R \in L_1^n(\Omega, \mathcal{F}, P)$.

The vector $z \in \mathbb{R}^n$ represents our asset allocation, with each component z_j equal to the fraction of capital invested in asset j. The set Z of feasible portfolios is given below $Z = \{z \in \mathbb{R}^n : \sum_{j=1}^n z_j \leq 1,\ z_j \geq 0,\ j = 1..n\}$.

In the analysis that follows, we will require only that Z is a convex, compact set in $\mathbb{R}^n$. So for example, one could limit the possible exposure of some assets by adding additional upper bounds on the investments in asset j, or on groups of assets. All these sets repesent closed convex sets in $\mathbb{R}^n$, so can be used. The total return of the portfolio at the end of the investment is $R^T z = \sum_{j=1}^n R^j z_j$.

The portfolio problem is to find an optimal way to invest the initial capital among the n assets, in the face of uncertainty about the returns R. This is usually approached by optimizing some objective function of the total return, over the set of feasible portfolios. The general portfolio problem, with objective function ρ is given by

$$\min_{z \in Z} \rho(R^T z) \tag{1.1}$$

We will examine the portfolio optimization problems where the objective function takes the form

$$\rho(R^T z) = -E[R^T z] + \gamma r[R^T z]. \tag{1.2}$$

This is the mean-risk approach introduced by Markowitz in his paper "Portfolio Selection" [18]. The term $r[R^T z]$ is a measure of the uncertainty of the portfolio return. In his paper, Markowitz used the variance of returns as the risk measure. The non-negative parameter γ represents our tolerance for risk. If $\gamma = 0$, then the problem reduces to a standard problem of maximizing the mean, and we are more tolerant of risk. The larger the value of γ, the more our tolerance for risk decreases. With this objective, we define the mean-risk portfolio problem

$$\min_{z \in Z} -E[R^T z] + \gamma r[R^T z]. \tag{1.3}$$

There has been extensive research on what risk functionals r should be used in the general mean-risk model. In the next section, we provide a more detailed interpretation of the meaning of a risk functionals r, given in the literature, and what properties a risk functional r should have. Examples of important risk functionals are given. We use the term “risk functional ” for the part $r[R^T z]$ representing the uncertainty of the return. The term “risk measure ” is used in the literature for the composite objective of the form (1.2).

1.3 Literature Survey

The term risk plays a pervasive role in much of the literature on financial and economic issues. Intuitively, it can be described as the chance of loss connected with a given action [6]. There have been many attempts to define and characterize risk in the literature, both for descriptive and prescriptive purposes. A detailed survey of these attempts can be found in [6]. For the purposes of financial risk, we use the definition of risk given in [41]. That is, risk is quantified on the basis of a random variable X. In this context, risk is interpreted as the potential loss or profit of a position. It can represent the future net worth of a portfolio, or the relative or absolute changes in an investment.

A risk measure is defined in [41] as a mapping from the space of random variables $\mathcal{X}$ representing outcomes to the real line.

Traditionally, the risk of a position was perceived as a dispersion in the values of the corresponding random variable. Since Markowitz [18] and Tobin [44], it was common to use the variance σ^2 and standard deviation σ to measure the dispersion of random variable X. The variance is defined as the average of the square of the deviations from the mean, $\sigma^2 = E[\ (X - E[X])^2\ ]$. The standard deviation is defined as the square root of the variance.

The variance and standard deviation have a number of nice properties. There are well established statistical methods for estimating these measures from data [41]. In particular, the mean-variance portfolio selection problem (1.3) can be reduced to a parametric quadratic programming problem, for which there are standard solution techniques.

An important criticism of the variance and standard deviation risk measures is that they penalize overperformance equally to underperformance. When the random variable X represents portfolio return, for example, returns above the mean are penalized. In keeping with the idea that risk should be a measure of some “negative occurance ”, the notion of *downside risk measures* was developed:

$$E[\ \max(c - X, 0)^k\ ] \tag{1.4}$$

The term c represents a target for which deviation below it is penalized. The number k is a measure of the relative impact of the deviations. Important examples include semideviation ($c = E[X], k = 1$) and semivariance ($c = E[X], k = 2$). Risk measures of the form (1.4) were examined in [9] for a fixed target value of c. There has been some disagreement over using a distribution-dependent target such as the mean for c. It has been argued by [17, 7] that risk is frequently associated with the failure to obtain a fixed target. To replace a set target c with a parameter (such as the mean) which changes from distribution to distribution, is not favourable to this model.
A variation of the downside risk measure is to take

$$[\ \ E[\ \max(c - X, 0)^k\ ]\ \ ]^{\frac{1}{k}} \tag{1.5}$$

Both (1.4) and (1.5) belong to one of the two larger classes of Stone’s risk measures (see [42] for a description of these classes).

In more recent research, there has been a push to develop axiomatic models of risk. Within this framework, there have been attempts to determine for which risk functionals the mean-risk models of the portfolio problem (1.3) will be consistent with these axioms. Two important axiomatic models are second-order stochastic dominance theory (SSD) and coherence axioms. A review of both is provided and we discuss the consistency of mean-risk models with these axioms.

Stochastic dominance [45] is based on an axiomatic model of risk-averse preferences [9]. It has its origins in majorization theory [14] for the discrete case, and was later extended to general distributions [13, 33]. It has been widely used in economics and finance. In the stochastic dominance approach, the random variables are compared by a pointwise comparison of some performance function, constructed from their distribution functions. The first performance function of a random variable X is defined as the right continuous cumulative distribution function

$$F_X(\eta) = P(X \leq \eta) \quad \forall \eta \in \mathbb{R}. \tag{1.6}$$

The weak relation of first-degree stochastic dominance (FSD) is defined by

$$X \succeq_{FSD} Y \Leftrightarrow F_X(\eta) \leq F_Y(\eta) \quad \forall \eta \in \mathbb{R}. \tag{1.7}$$

The second-order performance function is given by areas below the distribution function F

$$F_X^2(\eta) = \int_{-\infty}^{\eta} F_X(\xi) d\xi, \quad \eta \in \mathbb{R}. \tag{1.8}$$

A random variable X stochastically dominates random variable Y in the second order, if $F_X^2(\eta) \leq F_Y^2(\eta)$ for all $\eta \in \mathbb{R}$. We write $X \succeq_{SSD} Y$ for second-order stochastic dominance.

Strong FSD and SSD relations correspond to a strict inequality holding for at least one $\eta \in \mathbb{R}$. For decision making, the second order stochastic dominance relation is most important. The SSD relation is consistent with risk-averse preference models that prefer larger outcomes. A risk-averse investor's preferences can be described by a concave nondecreasing utility function $u : \mathbb{R} \to \mathbb{R}$. If $X \succeq_{SSD} Y$, then we have

$$E[u(X)] \geq E[u(Y)], \tag{1.9}$$

for all non-decreasing concave utility functions u. Thus a risk-averse investor would prefer position X over Y.

The consistency of mean-risk models with the second-order stochastic dominance relation has been examined in [9, 22, 23, 24, 25]. A mean-risk model is *SSD-consistent* if there exists a constant γ such that for all X,Y

$$X \succeq_{SSD} Y \;\Rightarrow$$
$$E[X] \geq E[Y] \; and \; E[X] - \gamma r(X) \geq E[Y] - \gamma r(Y)$$

In [25], the mean-risk model with the risk functional defined as semivariance (c̃orresponding to $k = 2$ in (1.4)) was found to be SSD-consistent, but the constant γ depends on the problem instance. This result was generalized by [9] to all mean-risk models with risk functionals of the form (1.4) for which $\gamma \geq 1$. In [22], the mean-risk model with r defined as the absolute semideviation (1.11) with $\gamma = 1$ is found to be SSD consistent. The mean-risk model with risk defined as deviation from quantile α with $\gamma = 1$ is also found to be SSD consistent. In [47], the mean-risk model with Gini's mean absolute difference was found to be SSD-consistent. In the case of discrete random returns, the mean-risk models can be formulated as linear programming problems, and the mean-risk efficient frontier calculated using fast parametric simplex method.

The coherence axioms are a more recent development, introduced in 1999 by Artzner et al in [2]. Denote by $\mathcal{X}$ the space of all uncertain outcomes. In the context of the portfolio problem, $X = R^T z$ and $\mathcal{X} = L_1(\Omega, F, P)$.

Definition 1. *A coherent measure of risk is a functional $\rho : \mathcal{X} \rightarrow \mathbb{R}$ which satisfies the following four axioms*:

A_1 *Convexity* : $\rho(\alpha X + (1-\alpha)Y) \leq \alpha\rho(X) + (1-\alpha)\rho(Y), \; \forall X, Y \in \mathcal{Z}$ *and* $\alpha \in [0, 1]$*;*

A_2 *Monotonicity* : *If* $X, Y \in \mathcal{X}$, *and* $X(\omega) \leq Y(\omega)$ *for all* $\omega \in \Omega$, *then* $\rho(X) \geq \rho(Y)$*;*

A_3 *Translation Property* : *If* $a \in R$ *and* $X \in \mathcal{X}$, *then* $\rho(X + a) = \rho(X) - a$*;*

A_4 *Positive Homogeneity*: *If* $t \geq 0$ *and* $X \in \mathcal{X}$*, then* $\rho(tX) = t\rho(X)$*.*

Important examples of coherent risk measures ρ are obtained from certain mean-risk models of the form

$$\rho(X) = -E[X] + \gamma r[X] \tag{1.10}$$

with scalar parameter $\gamma \geq 0$ and risk functional $r : X \to \mathbb{R}$ representing the variability of the return. In particular, we may set $r[X]$ to be the semideviation measure of order $p \geq 1$

$$r[X] = E[\ (E[X] - X)_+^p\ ]^{\frac{1}{p}} \tag{1.11}$$

or the weighted mean-deviation from quantile

$$r_\alpha[X] = \min_{\eta \in \mathbb{R}} E[\ \max(\frac{1-\alpha}{\alpha}(\eta - X),\ (X - \eta)\ ],\quad \alpha \in (0,1). \tag{1.12}$$

It is well known that the optimal η in the above problem is any $\alpha-$quantile of X. In both cases, when $\gamma \in [0,1]$, the resulting mean-risk model is coherent [35].

Coherent risk measures have a very important representation theorem. Under fairly mild conditions, when X represents final net worth, a coherent risk measure ρ can be represented as the supremum of the expected negative value of X over a set A of probability measures:

$$\rho(X) = \sup_{\mu \in A} E_\mu[-X]. \tag{1.13}$$

The mean-risk models with the variance and standard deviation risk functionals are not coherent [4]. Mean-risk models which are coherent include the mean-semideviation and mean-deviation from quantile models [35]. The analysis of the construction of quantile risk functionals is interesting in the context of coherent risk measures. We provide a brief history of it. We begin with the VaR risk measure.

The value at risk measure (VaR), was introduced by JP Morgan Chase in 1994. At a given probability level $\alpha \in (0,1]$, and random variable X representing the loss of a position, VaR_α measures the minimum loss incurred in the α percent worst cases of a portfolio. In [4], VaR_α at probability level $\alpha \in (0,1]$ is defined by

$$\mathrm{VaR}_\alpha(X) = -x^\alpha \tag{1.14}$$

where the upper quantile x^α is defined as

$$x^\alpha = \sup\{x : P[X \leq x] \leq \alpha\}. \tag{1.15}$$

VaR concentrates on the upper tail of the loss distribution. It is useful to risk managers concerned with the frequency of a default or probability of a loss, and not neccesarily its size. It is used by financial institutions to determine how much capital to put aside to control risk exposure [5] and how much capital is required for backing up trading activities. An important property that VaR satisfies is the *Law Invariance*. It is given in [41]:

$$\textit{Law Invariance} : \text{If } P[X \leq t] = P[Y \leq t] \ \forall t \in \mathbb{R}, \text{ then } \ \rho[X] = \rho[Y]; \tag{1.16}$$

Law invariance states that if two random variables have identical distributions, then the risk measure on those variables takes the same value. This is important in industrial and financial applications, where VaR has to be estimated from empirical data. An overview of different methods for estimating VaR is given in [8].

It turns out that VaR is not coherent. It satisfies the last three axioms, but violates convexity. In [4], it is argued that violation of convexity is a serious flaw, as it discourages portfolio diversification, an intuitive protection against risk. Rockafellar and Uryasev [32] introduced a risk measure related to VaR_α, which is both law invariant and convex. The result was the expected shortfall, or the CVaR risk measure.

Expected shortfall at probability level α is the average loss in the worst 100α percent cases [41, 4]. It is a measure of how much one can lose on average in states beyond VaR_α. The rigorous definition of ES_α is given in [4] by

$$\mathrm{ES}_\alpha[\mathrm{X}] = -\frac{1}{\alpha}(\ \mathrm{E}[\mathrm{X}1_{\{\mathrm{X}\leq \mathrm{x}^\alpha\}}]\ -\ \mathrm{x}^\alpha(\mathrm{P}[\mathrm{X}\leq \mathrm{x}^\alpha] - \alpha)\ ). \tag{1.17}$$

Here, 1_A is the indicator function on the set A, defined by

$$1_a = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \neq A \end{cases} \tag{1.18}$$

and x^α is the upper quantile given in (1.15). The risk measure ES_α is both law invariant and coherent. There are more intuitive representations of ES_α. If the generalized inverse function of $F(x)$ is introduced,

$$F^{-1}(p) = \inf\{x|F(x) \geq p)\} \tag{1.19}$$

it was shown in [4] that ES_α can be expressed as the negative mean of $F^{-1}(p)$ in the interval $[0, \alpha]$.

$$\mathrm{ES}_\alpha[\mathrm{X}] = -\frac{1}{\alpha}\int_0^\alpha \mathrm{F}^{-1}(\mathrm{p})\mathrm{dp}. \tag{1.20}$$

The authors in [24] argued that this formulation allowed for easier analysis of its properties. In [21] it is represented as

$$\mathrm{ES}_\alpha(\mathrm{X}) = -\mathrm{E}[\mathrm{X}] + \mathrm{r}_\alpha[\mathrm{X}] \tag{1.21}$$

where $r_\alpha[X]$ is the weighted mean-deviation from quantile risk measure. This representation is often referred to as the *Average Value at Risk* measure, or *AVaR*. The function $r_\alpha[X]$ is defined in [24], for X representing gain, by

$$r_\alpha[X] = \min_{\eta\in\mathbb{R}} E[\max(\frac{1-\alpha}{\alpha}(\eta - X),\ (X-\eta)],\ \alpha \in (0,1). \tag{1.22}$$

The expected shortfall risk measure ES_α, or AVaR_α, plays an important role in the theory of law-invariant coherent risk measures. The following result is due to Kusuoka [16] :

Theorem 2. *If* $(\Omega, \mathcal{F}, P)$ *is nonatomic, for every lower semicontinuous law invariant coherent measure of risk* $\rho[.]$ *on* $L_\infty(\Omega, F, P)$*, there exists a set* $\mathcal{N}$ *of probability measures on* $[0,1]$ *such that*

$$\rho[X] = \sup_{v \in \mathcal{N}} \int_0^1 \mathrm{AVaR}_\alpha[\mathrm{X}]\mathrm{v}(\mathrm{d}\alpha). \tag{1.23}$$

Thus the expected shortfall risk measure is the building block for law-invariant coherent risk measures. This result does not hold in general for the discrete case, however, there are some classes of risk functionals for which Theorem 2 holds.

The measurement of risk of a position over many time periods is different from the one-period risk measures discussed so far. In a portfolio problem, for example, with the option to rebalance, information may become available at some interum time period. In [37], the authors argue that this information may alter an investors perception of risk from the previous investment period. They develop conditional risk mappings to reflect this perception. The authors in [3] also discuss issues associated with information becoming available during an investment period. The authors in [3, 29], and others, have tried to develop axioms similar to the one-period coherence axioms (1).

Chapter 2

The Abstract Risk-Averse Portfolio Problem

2.1 Formal Statement of the Risk-Averse Portfolio Problem

In the next two sections, we formulate the abstract risk-averse portfolio optimization problem. Optimality conditions and duality theory for the problem are provided and important examples involving mean-risk models are given.

The portfolio problem was described in section (1.2). We review now the concept of a risk measure and provide the space of outcomes on which it is defined. The formulation of the abstract risk-averse portfolio problem follows that given in [35] and specialized in [21].

An uncertain outcome is represented by a function $X : \Omega \to \mathbb{R}$. In what follows, X represents the profit of a position. For example, X could be the return of a portfolio. By a risk measure we mean a real-valued function $\rho(X)$, defined on the set of uncertain outcomes $\mathcal{X}$. We use for $\mathcal{X}$ the space given in [35, 40],

$$\mathcal{X} = L_p(\Omega, F, P), \; p \in [1, +\infty). \tag{2.1}$$

This space is important as many risk measures are defined in terms of p-th order moments of a random variable. In the context of the portfolio problem, $X = R^T z$ and $\mathcal{X} = L_1(\Omega, F, P)$. We will assume that the risk measures ρ are proper, that is, $\rho(X) > -\infty$ for all $X \in \mathcal{X}$ and that the domain

$$\mathrm{dom}(\rho) := \{X \in \mathcal{X} : \rho(X) < +\infty\} \tag{2.2}$$

is non-empty. The abstract risk-averse portfolio problem, with coherent objective function ρ is given below

$$\min_{z \in Z} \rho(R^T z) \tag{2.3}$$

We observe that the function is convex and finite-valued on $\mathbb{R}^n$. It is therefore continuous [35]. As the set Z is compact, the minimum of $\rho(R^T z)$ is attained in Z, and the problem has an optimal solution.

2.2 Optimality and Duality Theory

In order to develop optimality conditions for problem (2.3), we recall the representation theorem of convex risk measures, first proved in [2] and then generalized in [10] and [35]. The version here follows that given in [35].

As in the previous section, we let $\mathcal{X}$ be the space of $\mathcal{F}$ measurable functions with finite p^{th} order moment

$$\mathcal{X} = L_p(\Omega, \mathcal{F}, P), \quad p \in [1, +\infty). \tag{2.4}$$

The dual space associated with $\mathcal{X}$ is the space $\mathcal{X}^* = L_q(\Omega, \mathcal{F}, P)$ of linear functionals on $\mathcal{X}$, with $q \in (1, \infty]$, and $\frac{1}{p} + \frac{1}{q} = 1$. The scalar product of $X \in \mathcal{X}$ and $\mu \in \mathcal{X}^*$ is defined as

$$\langle \mu, X \rangle := \int_\Omega \mu(\omega) X(\omega) dP(\omega). \tag{2.5}$$

The tuple $(\mathcal{X}, \mathcal{X}^*, \langle \cdot, \ \cdot \rangle)$ defines a paired topological space, and it is within this framework that the main representation theorem for convex risk measures is presented.

In [30], the conjugate function $\rho^* : \mathcal{X}^* \to \overline{\mathbb{R}}$ of a convex risk function $\rho : \mathcal{X} \to \overline{\mathbb{R}}$ is

$$\rho^*(\mu) := \sup_{X \in \mathcal{X}} \{\langle \mu, X \rangle - \rho(X)\} \tag{2.6}$$

and the biconjugate function ρ^{**} is

$$\rho^{**}(X) := \sup_{\mu \in \mathcal{X}^*} \{\langle \mu, X \rangle - \rho^*(\mu)\}. \tag{2.7}$$

The Fenchel Moreau Theorem [30] states that if ρ is proper, convex and lower semicontinuous, then $\rho = \rho^{**}$. That is,

$$\rho(X) = \sup_{\mu \in \mathcal{X}^*} \{\langle \mu, X \rangle - \rho^*(\mu)\}. \tag{2.8}$$

It is proved in [30] that the conjugate function ρ^* will be proper. Conversly, if (2.8) holds for some proper function $\rho^* : \mathcal{Z}^* \to \overline{\mathbb{R}}$, then ρ is proper, convex and lower semicontinuous. It is easily seen that (2.8) can be equivalently written as

$$\rho(X) = \sup_{\mu \in U} \{\langle \mu, X \rangle - \rho^*(\mu)\} \tag{2.9}$$

where $U = \mathrm{dom}(\rho^*)$. If the risk measure in addition satifies one or more of the coherent risk measure axioms, then more structure can be imposed on the set U, and a more compact representation of ρ is possible. We use the representation theorem given in [40].

Theorem 3. *Suppose that $\rho : \mathcal{X} \to \overline{\mathbb{R}}$ is convex, proper and lower semicontinuous. Then representation* (2.9) *holds with $U := dom(\rho^*)$. Moreover, we have that*

1. Condition (A_2) holds iff every $\mu \in U$ is non-positive, i.e. $\mu(\omega) \leq 0, \ \forall \in \Omega$;

2. Condition (A_3) holds iff $\int_\Omega \mu dP = -1$ for every $\mu \in U$;

3. Condition (A_4) holds iff the following representation holds

$$\rho(X) = \sup_{\mu \in U} \langle \mu, X \rangle. \tag{2.10}$$

It follows that if ρ is a coherent risk measure, and is proper and lower-semi continuous, then the representation

$$\rho(X) = \sup_{\mu \in U} \langle \mu, X \rangle \tag{2.11}$$

holds, with U being a subset of the following set

$$\beta := \{\mu \in \mathcal{X}^* : \int_\Omega \mu(\omega) dP(\omega) = -1,\ \mu \leq 0\}. \tag{2.12}$$

Moreover, by positive homogeneity, the set $U = \partial\rho(0)$. Thus the representation of coherent risk measures can be seen to follow naturally from the theory of convex functions.

We return to the portfolio problem. In this context, $X = R^T z$ and $\rho(.)$ is a given coherent risk measure. Suppose we define $A = -U$. We have by the representation theorem that

$$\rho(R^T z) = -\inf_{\mu \in A} \langle \mu, R^T z\rangle \tag{2.13}$$

The mean-risk models with semideviations and deviations from quantile satisfy the assumptions of the theorem and enjoy the representation (2.13). Owing to the theorem, the portfolio optimization problem (2.3), with $X = R^T z$ and $R \in L_p(\Omega, F, P)$ can be written as an inf-max problem

$$-\max_{z \in Z} \inf_{\mu \in A} \langle \mu, R^T z\rangle \tag{2.14}$$

If the risk measure ρ is continuous then the set A is bounded. As it is convex and closed, it is weakly* compact. Therefore the "inf" operation can be replaced by the "min" operation. Moreover, due to the compacteness of Z and weak* compactness of A, the "min" and "max" operations can be interchanged, and we can prove the main optimality theorem [35].

Theorem 4. *Suppose ρ is a continuous coherent measure of risk. A point $\bar{z}$ is an optimal solution of problem (2.3) $\Leftrightarrow \exists$ a convex and weakly* closed set $\overline{A} \subset A$ such that for all probability measures $\mu \in \overline{A}$ the point $\bar{z}$ is also a solution of the problem*

$$\max_{z \in Z} \langle \mu, R^T z\rangle. \tag{2.15}$$

Furthermore, the set $\overline{A}$ is the set of solutions to the dual problem

$$\min_{\mu \in A} \max_{z \in Z} \langle \mu, R^T z \rangle. \tag{2.16}$$

Proof. Let $F(z, u)$ be the function defined on $Z \times A$ as follows

$$F(z, \mu) = \int_{\Omega} R^T(\omega) z \mu(\omega) P(d\omega) \tag{2.17}$$

The set Z is compact in $\mathbb{R}^n$, and the set A is weakly compact in L_q. The function is concave-convex and thus it has a saddle point $(\overline{z}, \overline{\mu})$ on $Z \times A$:

$$F(z, \overline{\mu}) \leq F(\overline{z}, \overline{\mu}) \leq F(\overline{z}, \mu), \ \forall z \in Z, \ \forall \mu \in A. \tag{2.18}$$

□

It follows that the optimal portfolio $\overline{z}$ optimizes the expected return with respect to the optimal probability measure $\overline{\mu}$. We shall call $\overline{\mu}$ the risk-adjusted probability measure. From the dual problem, it is seen that $\overline{\mu}$ is the worst possible measure in the set A.

From now on, we assume the probability space Ω is finite, with m elementary events $\omega_1, ..., \omega_m$ occuring with probabilities $p_1, ..., p_m$. The vector $p \in \mathbb{R}^m$ denotes the set of probabilities with coordinates $p_i, \ i = 1..m$.

The matrix R will denote the possible asset returns: r_{ji} denotes the return of asset j, in event i, where $j = 1..n$ and $i = 1...m$. With this notation, Rp denotes the vector of expected asset returns, $R^T z$ denotes the vector of portfolio returns, and $p^T R^T z$ is the expected portfolio return. The measure μ will be interpreted as a vector in $\mathbb{R}^m$. In this notation, the portfolio problem (2.3) can be written equivalently as

$$\max_{z \in Z} \min_{\mu \in A} \langle \mu, R^T z \rangle \tag{2.19}$$

with the dual problem

$$\min_{\mu \in A} \max_{z \in Z} \langle \mu, R^T z \rangle \tag{2.20}$$

This representation allows to view the portfolio problem as a matrix game, with payoff matrix R^T and strategies of players represented by the portfolio allocation z and measure μ. Finding optimal asset allocation $\overline{z}$ and optimal risk-adjusted probability measure $\overline{\mu}$ is equivalent to finding a saddle point of the game, restricted to sets Z and A. In the next two sections, we show several important mean-risk models that can be formulated as linear programs in the discrete case, and we construct the risk adjusted probability measure $\overline{\mu}$ for these cases.

2.3 The Mean-Semideviation Model

The absolute semideviation risk measure r of a random variable X is defined as

$$\sigma[X] = E \max(E[X] - X, 0). \tag{2.21}$$

The corresponding mean-risk model in this case takes the form

$$\rho[X] = -E[X] + \gamma \sigma[X] \tag{2.22}$$

It was proved in [22] that $\rho(X)$ is consistent with second order stochastic dominance for $\gamma \in [0, 1]$. In [35], it was proved that $\rho(X)$ is coherent if $\gamma \in [0, 1]$. For discrete distributions, we can identify X with a vector in $\mathbb{R}^n$ and write

$$\rho[X] = -\langle p, X \rangle + \gamma \sum_{i=1}^{m} p_i \max(\langle p, X \rangle - x_i, 0), \tag{2.23}$$

where x_i denotes the ith outcome of random variable X, and p_i is its probability. The portfolio problem, with $X = R^T z$, becomes

$$\min_{z \in Z} -\langle p, R^T z \rangle + \gamma \sum_{i=1}^{m} p_i \max(\langle p, R^T z \rangle - \langle r_i, z \rangle, 0), \tag{2.24}$$

where $r_i \in \mathbb{R}^n$ represents the vector of asset returns corresponding to outcome i. Using the representation theorem from the previous section, the portfolio problem can be represented as

$$\max_{z \in Z} \min_{\mu \in A} \langle \mu, R^T z \rangle \tag{2.25}$$

with the dual problem

$$\min_{\mu \in A} \max_{z \in Z} \langle \mu, R^T z \rangle \tag{2.26}$$

Here $A = -\partial\rho(0)$ is a subset of the set of probability measures. The set A has been described in [35]. In our notation, with discrete distributions, it takes the form

$$-\partial\rho(0) = \{\mu : \mu = (1 - \langle g, 1 \rangle)p + g,\ 0 \leq g \leq \gamma p\} \tag{2.27}$$

The set A in this case can also be determined through the process of finding the convex programming dual problem, as we now show.

Consider the portfolio problem given in (2.24). Denoting the shortfall $\left[\langle p, R^T z \rangle - \langle r_i, z \rangle\right]_+$ by s_i, we can write the problem as a convex programming problem [21]:

$$\begin{aligned} \text{Minimize } & -\langle p, R^T z \rangle + \gamma \langle p, s \rangle \\ \text{s.t. } & s_i \geq \langle p, R^T z \rangle - \langle r_i, z \rangle \\ \text{s.t. } & s \geq 0, \qquad z \in Z \end{aligned} \tag{2.28}$$

Associate Lagrange multipliers ξ with the constraints in (2.28). The Lagrangian function is

$$\begin{aligned} L(z, s, \xi) &= -\langle p, R^T z \rangle + \gamma \langle p, s \rangle + \sum_{i=1}^{m} \xi_i (\langle p, R^T z \rangle - \langle r_i, z \rangle - s_i) \\ &= (\langle \xi, 1 \rangle - 1)\langle p, R^T z \rangle - \langle \xi, R^T z \rangle + \langle \gamma p - \xi, s \rangle \end{aligned}$$

The dual function $L_D(\xi)$ is given by

$$L_D(\xi) = \inf_{z \in Z, s \geq 0} L(z, s, \xi) \tag{2.29}$$

By separating the term involving s from the terms involving $R^T z$, the dual function can be written as the sum of the optimal values of two problems:

$$L_D(\xi) = \min_{z \in Z} \{(\langle \xi, 1 \rangle - 1)\langle p, R^T z \rangle - \langle \xi, R^T z \rangle\} + \min_{s \geq 0} \{\langle \gamma p - \xi, s \rangle\} \tag{2.30}$$

Recall that the dual problem is defined by

$$\max_{\xi \geq 0} L_D(\xi) \tag{2.31}$$

In order to simplify the presentation of (2.30), observe that $L_D(\xi) = -\infty$ unless the following condition holds

$$\gamma p \geq \xi \tag{2.32}$$

In this case, the dual function reduces to

$$L_D(\xi) = \min_{z \in Z} \langle (\langle \xi, 1 \rangle - 1)p - \xi, \ R^T z \rangle \tag{2.33}$$

Let A' denote the set of elements

$$\{\mu : \mu = (1 - \langle \xi, 1 \rangle)p + \xi, \ \gamma p \geq \xi, \ \xi \geq 0\} \tag{2.34}$$

Then the dual problem becomes

$$\max_{\mu \in A'} \min_{z \in Z} \langle -\mu, R^T z \rangle \tag{2.35}$$

We show that μ is a probability measure. Note that $\langle \mu, 1 \rangle = 1$. Moreover, due to relation (2.32),

$$\mu \geq (1 - \gamma \langle p, 1 \rangle)p + \xi = (1 - \gamma)p + \xi \geq 0 \tag{2.36}$$

It follows that μ is a probability vector. By substituting $\xi = \gamma g$, we observe that μ is an element of the set A defined in (2.27). Thus there is a one-to-one correspondence between the feasible points μ in the dual problem (2.35) and the elements of the set A in (2.27).

It follows that the convex programming dual problem (2.35) coincides with the game theoretic dual defined in (2.26). In this way, the following result has been proved.

Theorem 5. *Suppose $\rho(X) = -E(X) + \gamma\sigma_1(X)$ with $\gamma \in [0,1]$. A vector $\overline{z}$ and a measure $\overline{\mu}$ constitute a saddle point of game (2.25) $\Leftrightarrow$ the vector z is a solution of problem (2.28) and*

$$\overline{\mu} = (1 - \langle \overline{\xi}, 1 \rangle)p + \overline{\xi}, \tag{2.37}$$

where $\overline{\xi}$ are the Lagrange multipliers associated with constraints in (2.28).

It follows that we can obtain the risk adjusted probability measures by solving the convex programming problem (2.28), obtaining the Lagrange multipliers ξ, and applying the transformation in (2.37). When Z is a convex polyhedron, then linear programming methods can be employed.

2.4 The Mean -Weighted Deviation from Quantile Model

Consider the weighted deviation from α-quantile risk measure defined in (2.39):

$$r_\alpha[X] = \min_{\eta \in R} E\left[\ \max\left(\frac{1-\alpha}{\alpha}(\eta - X),\ (X - \eta)\right)\right],\ \alpha \in (0,1) \tag{2.38}$$

The corresponding mean-risk model in this case takes the form

$$\rho[X] = -E[X] + \gamma r_\alpha[X], \tag{2.39}$$

It was proved in [22] that $\rho(X)$ is consistent with second order stochastic dominance for $\gamma \in [0,1]$. In [35], it is proved that $\rho(X)$ is coherent for $\gamma \in [0,1]$. For discrete distributions, we identify X with a vector in $\mathbb{R}^m$ and write

$$\rho[X] = -\langle p, X\rangle + \gamma \min_{\eta\in\mathbb{R}} \sum_{i=1}^{m} p_i \max\left(\frac{1-\alpha}{\alpha}(\eta - x_i), x_i - \eta\right). \tag{2.40}$$

The portfolio problem, with $X = R^T z$ becomes

$$\min_{z\in Z} -\langle p, R^T z\rangle + \gamma \min_{\eta\in\mathbb{R}} \sum_{i=1}^{m} p_i \max\left(\frac{1-\alpha}{\alpha}(\eta - (R^T z)_i), (R^T z)_i - \eta\right) \tag{2.41}$$

Using the representation theorem from the previous section, the portfolio problem can be represented as

$$\min_{z\in Z} \max_{\mu\in A}\langle -\mu, R^T z\rangle \tag{2.42}$$

where $A = -\partial\rho(0)$ is a subset of the set of probability measures. The set A has been described in [35]. In our notation, with discrete distributions, it takes the form

$$A = \{\mu : \mu = (1-\gamma)p + \gamma g,\ 0 \le g_i \le \frac{p_i}{\alpha},\ i = 1..m,\ \ \langle g, 1\rangle = 1\}. \tag{2.43}$$

The set A in this case can also be determined through the process of finding the convex programming dual problem, as in the previous section.

Consider the portfolio problem (2.41). Denoting by u_i and v_i the excess $(x_i - \eta)$ and the shortfall $(\eta - x_i)$ respectively, the portfolio problem can be written as a convex programming problem (see [21]):

$$\begin{aligned}
&\text{Minimize } -\langle p, R^T z\rangle + \gamma \sum_{i=1}^{m} p_i \left(\frac{1-\alpha}{\alpha} v_i + u_i\right)\\
&\text{s.t. } \langle r_i, z\rangle - \eta = u_i - v_i, \qquad i = 1, .., m,\\
&\text{s.t. } u_i, v_i \ge 0, \quad i = 1, .., m,\\
&\text{s.t. } \eta \in \mathbb{R},\ z \in Z
\end{aligned} \tag{2.44}$$

Here, r_i denotes the return vector in the ith scenario. Associate Lagrange multipliers

ξ_i with the first set of constraints in (2.44). The Lagrangian function has the form

$$L(z, \eta, u, v, \xi) = -\langle p, R^T z\rangle + \gamma \sum_{i=1}^{m} p_i(\frac{1-\alpha}{\alpha} v_i + u_i)$$
$$+ \sum_{i=1}^{m} \xi(u_i - v_i - \langle r_i, z\rangle + \eta).$$

Collecting terms, we obtain

$$L(z, \eta, u, v, \xi) = -\langle p + \xi, R^T z\rangle - \eta\langle \xi, 1\rangle$$
$$+ \langle \frac{\gamma(1-\alpha)}{\alpha} p - \xi, v\rangle + \langle \lambda p - \xi, u\rangle.$$

The dual function $L_D(\xi)$ is given by

$$L_D(\xi) = \inf_{z\in Z, s\geq 0,\ \eta\in\mathbb{R},\ \mu, v\geq 0} L(z, \eta, u, v, \xi). \tag{2.45}$$

By separating the terms involving s, $R^T z$ and η, the dual function can be written as

$$L_D(\xi) = \min_{z\in Z} -\langle p+\xi, R^T z\rangle - \sup_{\eta\in R} \eta\langle \xi, 1\rangle + \inf_{v\geq 0} \gamma \left\langle \frac{1-\alpha}{\alpha})p - \xi,\ v \right\rangle + \inf_{u\geq 0} \langle \gamma p - \xi, u\rangle \tag{2.46}$$

The dual problem is given by

$$\max_{\xi\in\mathbb{R}^n} L_D(\xi). \tag{2.47}$$

In order to simplify the presentation of (2.46), observe that $L_D(\xi) = -\infty$ unless the following conditions hold:

$$\langle \xi, 1\rangle = 0,$$
$$-\gamma(\frac{1-\alpha}{\alpha})p \leq \xi \leq \gamma p.$$

In this case, the dual function reduces to

$$\min_{z\in Z} \langle -(p+\xi), R^T z\rangle). \tag{2.48}$$

Let A' denote the set of elements

$$\{\mu : \mu = p + \xi,\ \langle \xi, 1\rangle = 0\ ,\ -\gamma(\frac{1-\alpha}{\alpha})p \leq \xi \leq \gamma p\ \}, \tag{2.49}$$

Then the dual problem becomes

$$\max_{\mu \in A'} \min_{z \in Z} \langle -\mu, R^T z\rangle. \tag{2.50}$$

The conditions in A' and $\gamma \in [0,1]$ imply that μ is a probability vector. We have that $\langle \mu, 1\rangle = \langle p, 1\rangle + \langle \xi, 1\rangle = 1$, since second term is 0. To check that μ is non-negative, note that

$$\mu = p + \xi \geq p + \gamma p = (1+\gamma)p \geq 0 \tag{2.51}$$

By substituting $\xi = \gamma(g - p)$, we observe that μ is an element of the set A defined in (2.43). Thus there is a one-to-one correspondence between the feasible points ξ in the dual problem (2.50) and the elements of the set A in (2.43).

It follows that the convex programming dual problem (2.48) coincides with the game theoretic dual defined in (2.50). In this way, the following result has been proved.

Theorem 6. *Suppose $\rho(X) = -E[X] + \gamma r_\alpha(X)$ with $\gamma \in [0,1]$. A vector $\overline{z}$ and a vector $\overline{u}$ constitute a saddle point of the game (9) $\Leftrightarrow$ the vector z is a solution of problem (*2.41*) and*

$$\overline{\mu} = p + \overline{\xi}, \tag{2.52}$$

*where $\overline{\xi}$ are the Lagrange multipliers associated with problem (*2.44*).*

It follows that we can obtain the risk adjusted probability measures by solving the convex programming problem (2.41), obtaining the Lagrange multipliers ξ, and apply the transformation in (2.52). When Z is a convex polyhedron, then linear programming methods can be employed.

Chapter 3

Numerical Experiments (Part 1)

3.1 Objective and Setup

In this section, we find the optimal portfolios and compute the risk-adjusted probability measures for the mean-risk portfolio optimization problems based on the semideviation (problem (2.28)) and mean-deviation from quantile (problem(2.44)) measures of risk.

Each portfolio was drawn from a group of 200 assets taken from the $S\&P500$ index. Daily returns from the last 100 days of trading where taken as equally likely scenarios.

In each case, for risk aversion constant $\gamma = 0.5$, the cumulative distribution functions(CDF) for the optimal portfolio returns were constructed: one using original probability measures p, and the other using risk-adjusted probability measures μ. These CDF's were plotted against each other.

Separately, a market portfolio, with each asset having equal weight, was constructed. The CDF's for both the original probability measures and the risk-adjusted probability measures were constructed for this portfolio, and plotted against each other.

The shape of the cumulative distribution function provides a pictorial description of the behaviour of the portfolio. For example, if the curve takes larger values at negative returns, then the likelihood of poor portfolio perfomance is higher. So by plotting the CDF's for the original and risk-adjusted porbability measures, we are in a sense comparing the perspectives on the behaviour of the portfolio.

Recall that the risk-adjusted probability measure represents the worst possible measure in the set A for the matrix representation of the portfolio problem

$$\min_{z \in Z} \max_{\mu \in A} -\langle \mu, R^T z \rangle. \tag{3.1}$$

So a plot of the CDF of returns with respect to this risk-adjusted probability measure would in some sense reflect for that portfolio the worst possible behaviour for the returns. This is the curve used to reflect the perspective of a risk-averse investor.

If the CDF's for both measures are close together, then the portfolio is in keeping with the risk-averse investors' preferences. If the curves are far apart, the optimal portfolio does not reflect the concerns of a risk-averse investor. The former solution is considered robust. By constructing and comparing the CDF's for both the mean-risk optimal portfolios and the market portfolios, we can determine if the former method really does better reflect risk-averse investors preferences. The gaps in the former should be smaller than in the latter if this is true. Our hypothesis is that this is true.

3.2 Results

3.2.1 Mean-Semideviation Portfolio

The optimal portfolio for the mean-semideviation portfolio problem is presented in Table 3.1. In the table, the portfolio is heavily weighted towards three assets $(112, 138, 160)$, with the remaining capital dispersed more evenly among the other six assets. As discussed earlier, the moderate size of the risk-penalty constant $\gamma = 0.5$ corresponds to a moderately diversified portfolio.

The risk-adjusted probability measures for the mean-semideviation portfolio problem are presented in Table 3.2. The CDF of returns with respect to these measures is plotted against the CDF of returns with respect to the original probability measures in Figure 3.1. In the figure, the risk-adjusted CDF has slightly higher probability in the lower return values than the original CDF. This suggests a slightly more pessimistic outlook on the part of a risk-averse investor. The curves are somewhat close together, with a gap of value at most 0.1. This suggests a fairly robust portfolio. This corresponds to the Table 3.2, were the risk adjusted probability measures are fairly close in value to the original probability measures.

The risk-adjusted CDF is plotted against the original CDF in Figure 3.2 for the market portfolio. In this figure, the curves are also fairly close together, suggesting in this case, that even the market portfolio is somewhat robust.

So the hypothesis that the optimal portfolio is more robust than the market portfolio is not really supported for the mean-semideviation portfolio problem, with $\gamma = 0.5$. For larger γ coefficients, corresponding to a more risk-averse investor, the relationship may change.

Asset	Value
z_{15}	0.035
z_{33}	0.061
z_{73}	0.016
z_{99}	0.037
z_{112}	0.34
z_{138}	0.12
z_{160}	0.30
z_{164}	0.04
z_{178}	0.05

Table 3.1: The mean-semideviation optimal portfolio

2	0.0077	21	0.0127	40	0.0127	59	0.0079	78	0.0127	97	0.0127
3	0.0127	22	0.0127	41	0.0127	60	0.0127	79	0.0127	98	0.0077
4	0.0077	23	0.0077	42	0.0127	61	0.0077	80	0.0127	99	0.0077
5	0.0077	24	0.0077	43	0.01207	62	0.0127	81	0.0077	100	0.0077
6	0.0077	25	0.0106	44	0.0127	63	0.0077	82	0.0127	101	0.0127
7	0.0077	26	0.0077	45	0.0077	64	0.0077	83	0.0119		
8	0.0127	27	0.0127	46	0.0127	65	0.0127	84	0.0127		
9	0.0077	28	0.0077	47	0.0077	66	0.0126	85	0.0127		
10	0.0077	29	0.0127	48	0.0127	67	0.0077	86	0.0093		
11	0.0127	30	0.0127	49	0.0127	68	0.0127	87	0.0127		
12	0.0077	31	0.0127	50	0.0127	69	0.0077	88	0.0127		
13	0.0077	32	0.0077	51	0.0127	70	0.0127	89	0.0127		
14	0.0077	33	0.0077	52	0.0077	71	0.0127	90	0.0077		
15	0.0127	34	0.0077	53	0.0127	72	0.0077	91	0.0127		
16	0.0077	35	0.0077	54	0.0077	73	0.0077	92	0.0127		
17	0.0127	36	0.0077	55	0.0077	74	0.0077	93	0.0077		
18	0.0127	37	0.0077	56	0.0107	75	0.0077	94	0.0077		
19	0.0077	38	0.0077	57	0.0077	76	0.0077	95	0.0127		
20	0.0077	39	0.0077	58	0.0077	77	0.0077	96	0.0092		

Table 3.2: Risk-adjusted probability measures for the mean-semideviation optimal portfolio

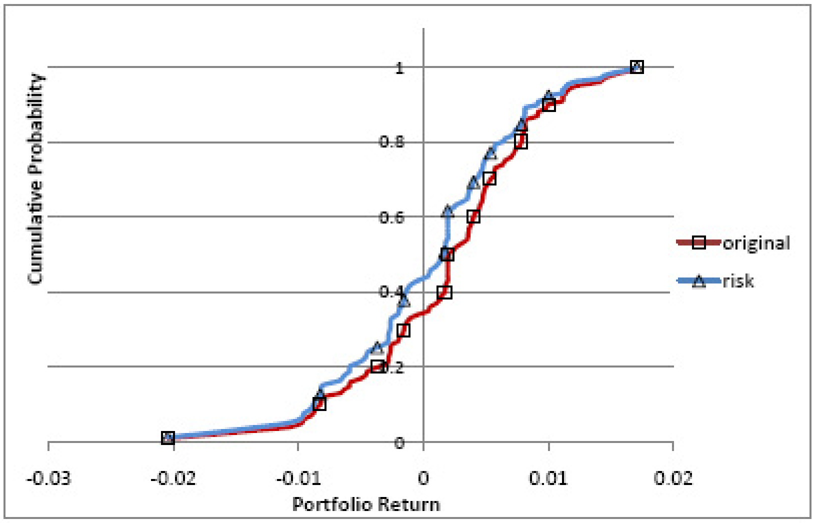

Figure 3.1: Cumulative distribution curves for the returns of the mean-semideviation optimal portfolio.

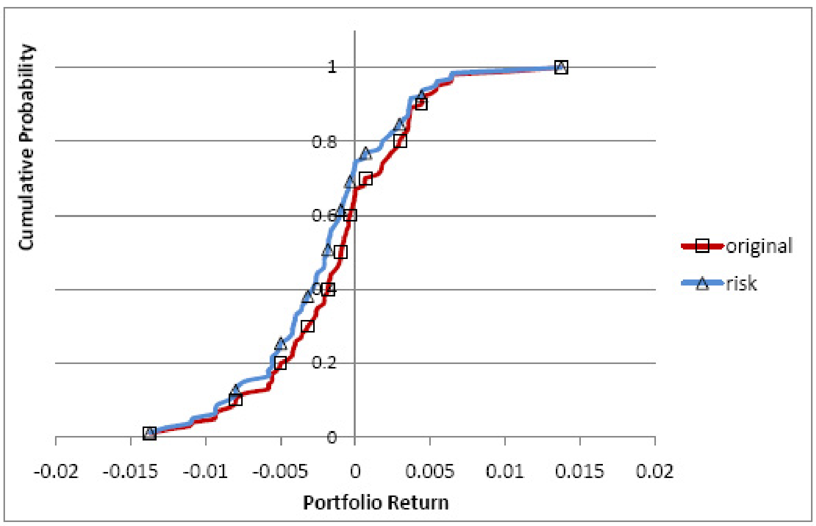

Figure 3.2: Cumulative distribution curves of market portfolio for the semideviation risk function.

3.2.2 Mean-Weighted Deviation From Quantile

The optimal portfolio for the mean-weighted deviation from quantile portfolio problem is presented in Table 7.3.

The risk-adjusted probability measures for the optimal mean-weighted deviation from quantile portfolio problem are presented in Table 3.4. The risk-adjusted CDF is plotted against the original-CDF in Figure 3.4 for the optimal portfolio. In the figure, the curves are very close together. This suggests that the optimal portfolio is strongly robust. This corresponds with Table 3.4, were the risk-adjusted probability measures are close in value to the orignal probability measures.

The risk-adjusted CDF is plotted against the original-CDF in Figure 3.3 for the market portfolio. In this figure, the curves diverge substantially. The risk-adjusted measures predict a much more pessimistic outcome for the returns than the original probability measures. This can be seen by the high value of the risk-adjusted CDF in the negative half of returns, compared to the lower original CDF values in this interval.

The hypothesis that the optimal portfolio is more robust than the market portfolio is strongly supported for the mean-weighted deviation from quantile portfolio problem. The variation in the two curves can be explained by examining the risk functional. In the weighted deviation from quantile risk functional, returns below the α-quantile are penalized with coefficient $\frac{1-\alpha}{\alpha}$. For the five percent quantile, this coeeficient takes the value 19, which is very high. For a market portfolio, with equally weighted assets, and no effort to avoid this left tail, the penalty is applied with much higher frequency.

asset	value	asset	value	asset	value	asset	value
z2	0.055	z5	0.062	z8	0.007	z9	0.011
z12	0.007	z27	0.037	z30	0.023	z33	0.008
z37	0.011	z42	0.010	z44	0.014	z46	0.019
z58	0.066	z67	0.0542	z73	0.003	z84	0.005
z94	0.030	z102	0.034	z104	0.015	z105	0.015
z112	0.057	z119	0.102	z120	0.026	z127	0.004
z133	0.068	z136	0.054	z138	0.026	z141	0.034
z160	0.067	z177	0.005	z182	0.018	z184	0.018
z191	0.026	z193	0.01				

Table 3.3: The mean-deviation from quantile optimal portfolio

2	0.005	27	0.022	52	0.005	77	0.005
3	0.005	28	0.018	53	0.050	78	0.009
4	0.005	29	0.005	54	0.007	79	0.017
5	0.005	30	0.005	55	0.005	80	0.006
6	0.005	31	0.005	56	0.005	81	0.005
7	0.010	32	0.005	57	0.005	82	0.028
8	0.005	33	0.015	58	0.005	83	0.005
9	0.005	34	0.005	59	0.005	84	0.005
10	0.005	35	0.005	60	0.005	85	0.028
11	0.005	36	0.005	61	0.005	86	0.005
12	0.005	37	0.005	62	0.026	87	0.023
13	0.005	38	0.005	63	0.005	88	0.005
14	0.009	39	0.005	64	0.008	89	0.021
15	0.042	40	0.005	65	0.019	90	0.005
16	0.005	41	0.005	66	0.009	91	0.048
17	0.02	42	0.011	67	0.005	92	0.049
18	0.02	43	0.005	68	0.005	93	0.005
19	0.005	44	0.005	69	0.005	94	0.005
20	0.005	45	0.005	70	0.005	95	0.01
21	0.005	46	0.005	71	0.011	96	0.005
22	0.037	47	0.008	72	0.018	97	0.007
23	0.005	48	0.005	73	0.005	98	0.005
24	0.012	49	0.005	74	0.005	99	0.007
25	0.018	50	0.005	75	0.005	100	0.005
26	0.005	51	0.027	76	0.005	101	0.005

Table 3.4: Risk-adjusted probability measures for the mean-deviation from quantile optimal portfolio

.

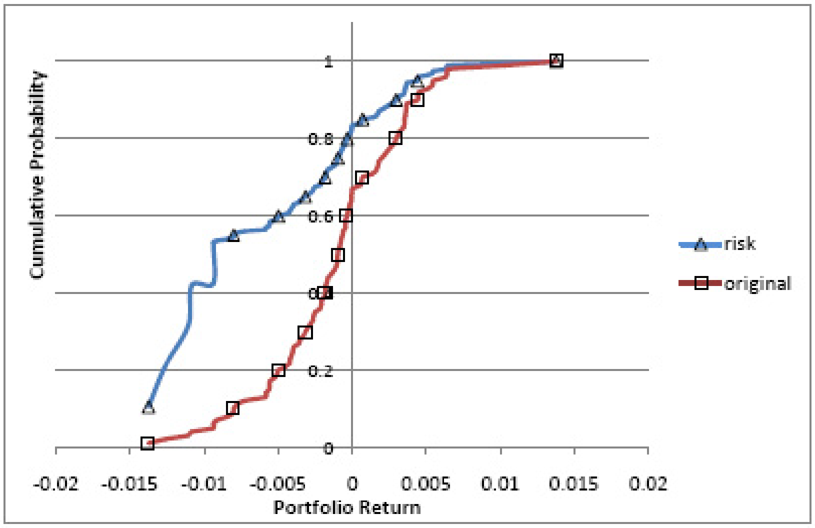

Figure 3.3: Cumulative distribution curves of the market portfolio for the deviation from quantile risk function.

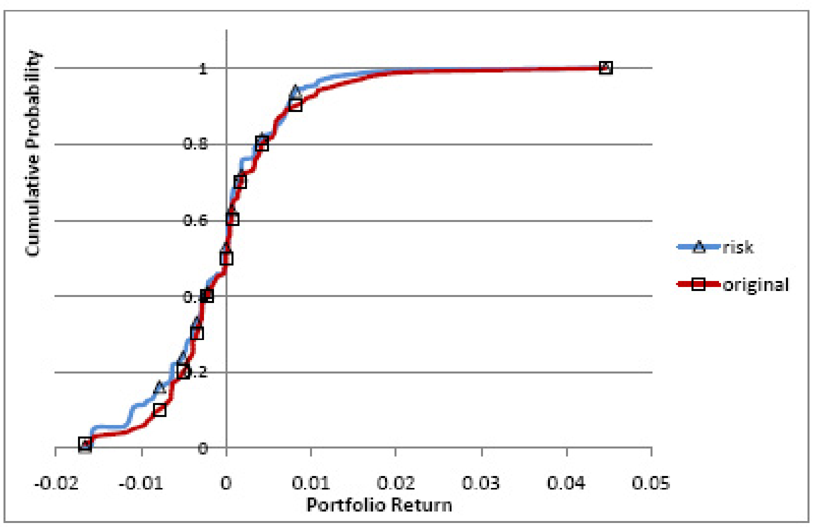

Figure 3.4: Cumulative distribution curves for the returns of the mean-deviation from quantile optimal portfolio.

Chapter 4

The Two- Stage Portfolio Problem

4.1 Formulation of the Standard Two-Stage Portfolio Problem

In the first part of this dissertation, we introduced the risk-averse approach to optimizing a portfolio of n assets over one investment time period. In this approach, we formulated the following optimization problem

$$\min_{z \in Z} \rho(R^T z), \tag{4.1}$$

where ρ was a coherent risk measure. The coherence of ρ allowed us to rewrite the problem as a matrix game and to derive new risk-adjusted probability measures. The risk-averse approach was argued to reflect the attitudes of a risk-averse investor.

In this section, our objective is to formulate an analogous risk-averse approach to optimizing a portfolio over two time periods. In particular, we are interested in the case where the option exists to rebalance the portfolio in between the two time periods. With this in mind, we formulate the conditional risk mapping approach for optimizing the portfolio and argue for its benefits. The resulting two-stage portfolio optimization problem is called a two-stage risk-averse portfolio problem.

We begin with a review of the two-stage portfolio problem with rebalancing. Consider a collection of n assets in which investment decisions are to be made in two consecutive time periods. The return of the assets in each stage are assumed to be n - dimensional integrable random variables, R^t on some probability space, with R^t_j the return of asset j in stage t, for $t \in \{1, 2\}$.

Our asset allocations in the first and second stage are denoted by the n-dimensional vectors z^t, with components z_j^t representing the amount of capital invested in asset j during stage t, for $t \in \{1, 2\}$. The end portfolio value in stage t is given by $(\xi^t)^T z^t$, where

$$\xi^t = 1 + R^t \tag{4.2}$$

In the portfolio problem with rebalancing, the capital at the end of the first stage, given by $(\xi^1)^T z^1$ is reallocated among the assets, prior to observing the second stage return outcomes.

In what follows, we consider the portfolio problem for the discrete case, where the vector random variables ξ^1 and ξ^2 have a finite number of realizations. In this case, we can visualize the possible sequence of outcomes $\xi = (\xi^1, \xi^2)$ by a scenario tree

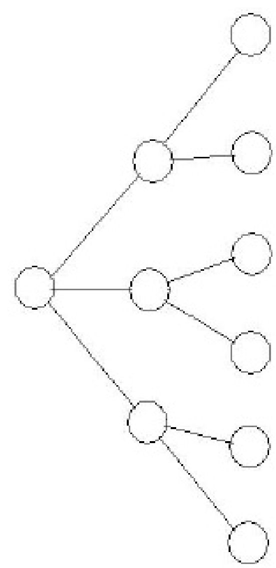

Figure 4.1: Scenario tree.

The nodes in level one and two represent the possible realizations of ξ^1 and ξ^2, while the root node at level 0 represents the beginning of the process. Each node i in level one is connected to a set of *children nodes*, C_i in level two, representing the possible outcomes of ξ^2 following the first stage outcome. The root node connects to all nodes in level one, and a path from the root node to an end node represents a sequence of possible outcomes for (ξ^1, ξ^2).

We can associate with the root node a probability vector $p^1 \in R^{m_1}$, with p^1_i the probability of outcome i occuring in the first stage. Similarly, we can associate with each node i in the first level, the probability vector $p^2_i \in R^{m_2}$, with p^2_{il} the probability of moving to node l in level two from this node.

Note that z^2 becomes an $n \times m_1$ matrix, with entry z^2_{ji} representing the second stage asset allocation, given that outcome i occured in the first stage. We can formulate explicitly the form of the sets Z_{2i} for the case where the portfolio is rebalanced after the first stage, subject to trading costs:

$$Z_{2i} = \{z^2_i : \sum_{j=1}^{n} \xi^1_j z^1_j - \kappa \sum_{j=1}^{n} |z^2_{ji} - \xi^1_j z^1_j| \geq \sum_{j=1}^{n} z^2_{ji}, \quad z^2_i \geq 0\}, \quad \kappa \in [0,1] \tag{4.3}$$

Here, the total trading costs are given by term $\kappa \sum_{j=1}^{n} |z^2_{ji} - \xi^1_j z^1_j|$. In the literature, these are calculated for each asset j as a non-negative fraction of the amount traded. The total amount invested in the second stage must not exceed the first stage end portfolio minus the total trading costs. This is given by the main constraint in the set.

The two stage portfolio problem with rebalancing fits into the class of stochastic programming problems with recourse, described in [27] and [40]. That is, we have a decision problem over m time periods, $m \geq 2$, where at least one decision is preceded by an observation. The decision process for the two-stage problem can be represented by the following diagram.

$$1.\ \textit{Decision}\ z^1 \rightarrow 2.\ \textit{Observation}\ \xi^1$$

$$3.\ \textit{Decision}\ z^2 \rightarrow 4.\ \textit{Observation}\ \xi^2$$

A standard approach described in [27] and [40] to solve this problem is to construct a linear two-stage problem

$$\begin{aligned} &\text{Minimize } c^T z^1 + E[Q(z^1)] \\ &\text{s.t. } Az^1 = b, \quad z^1 \geq 0 \end{aligned} \tag{4.4}$$

where $Q_i(z^1)$ is the optimal value of the ith second stage problem

$$\begin{aligned} &\text{Minimize } q_i^T z_i^2 \\ &\text{s.t. } T_i z^1 + W_i z^2 = h_i, \quad z_i^2 \geq 0 \end{aligned} \tag{4.5}$$

In these formulations, some or all of the vectors in the 4-tuple (q, W, T, h) are random. In many cases, the vector q_i represents the conditional expectation of a random function, $-E_{p_i^2}[\xi_i^2]$. The first and second stage feasible sets are closed and convex in $\mathbb{R}^n$. In the case of the portfolio problem, we can replace the sets with Z_1 and Z_{2i}. The difficulty arises that the set Z_{2i} is not defined by an inequality involving a linear function. It can be proved however that (4.5) has the same solution when Z_{2i} is replaced by the following convex set

$$\begin{aligned} \overline{Z}_{2i} = \{(z_i^2,\ u_i,\ v_i) : \sum_{j=1}^{n} \xi_j^1 z_j^1 - \kappa \sum_{j=1}^{n} (u_{ji} + v_{ji}) \geq \sum_{j=1}^{n} z_{ji}^2 \\ u_{ji} - v_{ji} = \xi_j^1 z_j^1 - z_{ji}^2, \quad z_i^2 \geq 0\ u \geq 0,\ v \geq 0\} \quad \kappa \in [0,1] \end{aligned} \tag{4.6}$$

With the set Z_{2i} replaced by $\overline{Z}_{2i}$, the linear two-stage problem takes the form

$$\min_{z^1 \in Z_1} \{c^T z^1 + E[Q(z^1)]\} \tag{4.7}$$

with $Q_i(z^1)$ is the optimal value of

$$\min_{z_i^2 \in Z_{2i}} c_{2i}^T z_i^2 + (-E_{p_i^2}[\xi_i^2])^T z_i^2 \tag{4.8}$$

We can visualize the linear two-stage approach (4.7) - (4.8) by considering what happens after observing ξ^1. At this point, outcome i has occurred, and this is represented by node i in the scenario tree. At node i, the optimization problem is to maximize the expected value of the portfolio at the end of the second stage, where the expectation is taken over the children nodes C_i. That is, the conditional expectation of $\xi^2 z^2$, given that we are at node i, must be optimized. The first stage problem is to optimize the portfolio over all possible outcomes i.

4.2 Formulation of the Risk-Averse Two-Stage Portfolio Problem

The conditional risk mapping approach to portfolio optimization builds on this method. That is, the decision-making for stage two is based on the restricted set of possible outcomes given we are at node i. However, we may choose, instead of conditional expectation, any coherent risk measure. For example, we may use conditional mean-semideviation

$$\rho_{2i}(Z) = -E_{p_i^2}[Z] + \gamma_i E_{p_i^2} \max((Z - E_{p_i^2}[Z]),\ 0),\ \ \gamma_i \in [0,1] \tag{4.9}$$

or conditional mean-deviation from quantile

$$\rho_{2i}(Z) = -E_{p_i^2}[Z] + \gamma_i E_{p_i^2} \max(\frac{1-\alpha}{\alpha}(Z - \eta),\ (\eta - Z)),\ \ \gamma_i \in (0,1) \tag{4.10}$$

Here, we set $Z = \xi_i^2 z_i^2$ for notational convenience. Suppose we denote by $Q_i(z^1)$ the optimal solution to the optimization problem at node i.

$$Q_i(z^1) = \min_{z_i^2 \in Z_{2i}} c_{2i}^T z_i^2 + \rho_{2i}(\xi_i^2 z_i^2) \tag{4.11}$$

As in the standard stochastic programming approach, we optimize some composite function ρ_1 over all 2nd stage optimal solutions. That is, we formulate the first stage problem

$$\min_{z^1 \in Z_1} c^T z^1 + \rho_1(-Q(z^1)) \tag{4.12}$$

Here, ρ_1 is a coherent risk measure, and $Q(z^1)$ is the random variable taking the value $Q_i(z^1)$ with probability p_i^1. We call problem (4.12) the risk-averse two-stage portfolio optimization problem. If we choose $\rho_1 = -E[.]$, then we obtain model (4.4).

We provide briefly some explanantion for why the composition of ρ_1 is taken with respect to $-Q(.)$. Recall the monotonicity condition

$$\text{If } X, Y \in \mathcal{Z}, \text{ and } X(\omega) \leq Y(\omega) \text{ for all } \omega \in \Omega, \text{ then } \rho(X) \geq \rho(Y)$$

We note that the risk measure ρ is actually negatively monotone. That is, ρ decreases in value as X increases in value. Thus, if $X = Q(.)$ represents a convex function, which is non decreasing, then the composition $\phi = \rho(Q(.))$ would result in a non-increasing and concave function. As we are interested in coherent mean-risk models, we require this composition to be nondecreasing and convex. Hence we take $-Q(.)$. The following proposition formalizes this argument [40]:

Proposition 7. *Let $\mathcal{X}$ be an L_p space. If the mapping $Q : \mathbb{R}^n \to \mathcal{X}$ is convex and $\rho : \mathcal{X} \to \mathbb{R}$ satisfies conditions (A_1) and (A_2), then the composite function $\phi(.) := \rho(-Q(.))$ is convex.*

4.3 The Time Consistency Property

Consider now the function ρ_1 in problem (4.12). We can write $Q_i(z^1)$ in the following form

$$Q_i(z^1) = X_{2i} + \rho_{2i}(X_{3i}) \tag{4.13}$$

where $X_{2i} = c_2^T z_i^2$ and $X_{3i} = \xi_i^2 z_i^2$. Substituting this expression for $Q_i(z^1)$ and letting $X_1 = c^T z^1$, we can rewrite the objective function ρ in (4.12) as

$$\rho = X_1 + \rho_1(-X_2 - \rho_{2|1}(X_3)\ ) \tag{4.14}$$

Here, $\rho_{2|1}$ reflects the dependence of the second stage function ρ_2 on the first stage outcome. We have in (4.14) a nested formulation of coherent risk functions ρ_1 and $\rho_{2|1}$. This expression for ρ in (4.14) motivates the following definition [40]

Definition 8. *A risk measure ρ representable in the form (4.14) for ρ_1 and $\rho_{2|1}$ coherent risk functions, is called a time consistent risk measure.*

We can interpret the meaning of time consistency of a risk measure by considering the scenario tree framework in which the conditional risk mapping approach was formulated. Time consistency means that with every node of a scenario tree in the first stage, is associated a coherent risk measure applied to the children nodes in the next stage. Thus, the information obtained from the first stage outcome is used to restrict the second stage outcome space over which a second stage problem is to be optimized.

We can interpret the property of time consistency from the perspective of analyzing future risk. Suppose that we are at node i after the first investment period, and we have the opportunity to rebalance the portfolio. Suppose also that we have knowledge as an investor of the scenario tree. Future risk of a position is a function of the uncertainty of the future outcome. What time consistency property says is that the information avalable at node i informs our perception of future risk by reducing the outcome space of end portfolio returns to the children nodes C_i.

We provide an algebraic representation of the property of time consistency for the standard 2-stage stochastic programming problem and extend it to the general 2-stage stochastic programming problem

$$\min_{z^1 \in Z_1} \rho_1(-Q(z^1)). \tag{4.15}$$

This representation proves useful for checking other approaches for this property.

Consider the standard 2-stage stochastic programming problem

$$\min_{z^1 \in Z_1} c^T z^1 + E_{p^1}[\inf_{z^2 \in Z_2} -E_{p^2}[\xi^2 z^2]]. \tag{4.16}$$

Our objective is to formulate problem (4.16) as one large linear programming problem. Here z^2 is a function of the first stage event. Using the interchangability principle [40], we obtain that (4.16) is equivalent to

$$\min_{z^1 \in Z_1,\ z^2 \in Z_2} c^T z^1 - E_{p^1}[E_{p^2}[\xi^2 z^2]] \tag{4.17}$$

In the discrete case, expanding the expectation function with respect to p^1 and p^2, we obtain

$$\min_{z^1 \in Z_1,\ z_i^2 \in Z_{2i}} c^T z^1 + \sum_{i=1}^{m_1} p_i^1 (\sum_{k=1}^{m_2} p_{ik}^2 [\xi_{ik}^2 z_i^2]) \tag{4.18}$$

To simplify this expression, let the variable r_{ik}^2 denote the term $\xi_{ik}^2 z_i^2$ and let r_i^1 denote the term $\sum_{k=1}^{m_2} p_{ik}^2 r_{ik}^2$. The problem becomes

$$\begin{aligned}
\text{Minimize } c^T z^1 \ &- \ \sum_i p_i^1 r_i^1 \\
\text{s.t. } r_i^1 &= \sum_{k=1}^{m_2} p_{ik}^2 r_{ik}^2 \quad i = 1..m \\
\text{s.t. } r_{ik}^2 &= \sum_{j=1}^{n} \xi_{jik}^2 z_{ji}^2, \quad i = 1..m,\ k = 1..m_2 \\
\text{s.t. } z^1 &\in Z_1,\ z_i^2 \in Z_{2i},\ r_{ik}^2 \geq 0\ \forall i, k
\end{aligned}$$

Note that if the first stage variables are fixed, then the second stage constraints corresponding to pair (i, k) are separate with respect to outcome i. That is, for each outcome i, we can obtain a set of second stage constraints which are a function of i only. This is equivalent to solving m_1 separate problems, and we call this the decomposition property. We can show in a similar manner for a general time consistent risk measure ρ_1 in the problem

$$\min_{z^1 \in Z_1} \rho_1(-Q(z^1)) \tag{4.19}$$

that the decomposition property holds. Thus we can determine if the risk measure in problem (4.19) is time consistent by formulating a large linear programming problem and checking for the decomposition property. We use this to test a common approach, the aggregate method to portfolio optimization, for time consistency.

In the aggregate model for portfolio optimization, the idea is to optimize in the first stage a function of all second stage outcomes. The aggregate model formulation is

$$\min_{z^1 \in Z_1,\ z^2 \in Z_2} \rho_1\Big(\sum_j \xi^2_{jik} z^2_{ji}\Big) \tag{4.20}$$

where $\rho_1(.)$ is some risk functional. This model is more intutively appealing than the conditional risk mapping approach in many cases. For example, we mentioned earlier that the functions ρ_1 and $\rho_{2|1}$ could both be mean-semideviation risk measures. But the question then arises about what the composition of two such functions really measures. In the aggregate method, we optimize one function of all 2nd stage outcomes, and there is more clear understanding of what is being measured. But is the aggregate model always time consistent? And if so, can we find a composition of two coherent risk measures which would produce the same value? We prove that the answer to the first question is no,for at least one case of ρ_1, using as an example, the model with a mean-semideviation objective.

Consider the aggregate model for the semideviation risk measure, given by

$$\begin{aligned} \min_{z^1 \in Z_1,\ z^2 \in Z_2} & -\sum_i p^1_i \sum_k p^2_{ik}(\xi^2_{ik}) \ + \\ & \gamma \sum_i p^1_i \sum_k p^2_{ik} \ \max\Big(\sum_l p^1_l \sum_t p^2_{lt}(\xi^2_{lt}) - \xi^2_{ik},\ 0\Big) \end{aligned} \tag{4.21}$$

We expand the aggregate model into a linear programming problem by letting r^2_{ik} denote the term $\max(\sum_l p^1_l \sum_t p^2_{lt}(\xi^2_{lt}) - \xi^2_{ik},\ 0)$, u_1 denote the term $\sum_i p^1_i \sum_k p^2_{ik}(\xi^2_{ik})$, and X^2_{ik} the term $\sum_j \xi^2_{jil} z^2_{ji}$. The aggregate model becomes

$$\text{Minimize } -u_1 \; + \; \gamma \sum_i p_i^1 \sum_k p_{ik}^2 r_{ik}^2$$

$$\text{s.t. } u_1 = \sum_i p_i^1 \sum_k p_{ik}^2 (X_{ik}^2)$$

$$\text{s.t. } r_{ik}^2 \geq u_1 - X_{ik}^2 \;\; i = 1..m, \;\; k = 1..m_2$$

$$\text{s.t. } z^1 \in Z_1, \;\; z^2 \in Z_2, \;\; r_i \geq 0 \;\forall i$$

We notice that in each of the second stage constraints, the mean u_1 is present. The mean u_1 requires the calculation of all second stage outcomes over all first stage outcomes i. Thus it is not separable into m_1 distinct second stage problems corresponding to outcomes i in the first stage. That is, the model does not satisfy the decomposition property, and so the risk function in the aggregate model is not time consistent.

Chapter 5

The Two-Stage Risk-Averse Portfolio Problem

5.1 Optimality and Duality Theory

In order to develop optimality conditions for the risk-averse portfolio problem (4.12) – (4.11), we recall the main representation theorem for general coherent risk measures, first proved in [2] and refined and generalized in a series of papers [10, 12, 31, 36, 35]. We use here a special case of the version given in [35].

Let $(\Omega, \mathcal{F}, P)$ denote a probability space, on which a space of $\mathcal{F}-$ measurable functions $\mathcal{X}$ is defined. As in previous sections, we will assume that $\mathcal{X} := L_p(\Omega, \mathcal{F}, P)$ for $p \in [1, +\infty)$. We recall the general composite risk mapping, given by

$$\phi(z) := \rho(-Q(z)) \tag{5.1}$$

where $\rho : \mathcal{X} \to \mathbb{R}$ is a risk measure and $Q : \mathbb{R}^n \to \mathcal{X}$ is a random function . We write $Q(z)[\omega]$ or $q(., \omega)$ for a particular outcome. Note that $Q(z)$ is an element of the space $L_p(\Omega, \mathcal{F}, P)$ and thus $q(z, .)$ is $\mathcal{F}$-measurable and finite valued. The mapping Q is said to be convex if the function $q(., \omega)$ is convex for every $\omega \in \Omega$. The two-stage stochastic programming problem can be formulated as

$$\min_{z \in Z} \rho(-Q(z)). \tag{5.2}$$

We will assume that the risk measure ρ satisfies the assumptions of convexity (A_1) and monotonicity (A_2). As the following proposition shows, these assumptions are enough to ensure continuity and subdifferentiability of ρ. See [40]:.

Proposition 9. *Let $\mathcal{X} := L_p(\Omega, \mathcal{F}, P)$ with $p \in [1, +\infty)$, and $\rho : \mathcal{X} \to \mathbb{R}$ be a real-valued risk measure satisfying conditions (A_1) and (A_2). Then ρ is continuous and subdifferentiable on $\mathcal{X}$.*

As Proposition 7 shows, the convexity (and hence subdifferentiability) of the composite risk measure is ensured if the random function Q is convex and the risk measure ρ satisfies the same assumptions as in Proposition 9

We make a note regarding this proposition. If in addition to the assumptions in Proposition 9 we assume that ρ satisfies translational equivariance (A_3), then we have the following representation theorem for ρ and its subdifferential [40].

Theorem 10. *Let $\mathcal{X} := L_p(\Omega, \mathcal{X}, P)$ and $Q : \mathbb{R}^n \to \mathcal{X}$ be a convex function. Let $\rho : \mathcal{X} \to \mathbb{R}$ be a convex function satisfying assumptions $(A_1) - (A_3)$. Then*

$$\begin{aligned} \rho(-Q(z)) &= \max_{v \in U} \{\langle v, -Q(z)\rangle - \rho^*(v)\}, \qquad U = dom(\rho^*) \\ &= \max_{\mu \in A} \{\langle \mu, Q(z)\rangle - \rho^*(-\mu)\}, \ \ A = -U. \end{aligned} \tag{5.3}$$

Furthermore, the subdifferential $\partial\rho(-Q(z))$ is given by

$$-\partial\rho(-Q(z)) = \arg\max_{\mu \in A} \{\langle \mu, Q(z)\rangle - \rho^*(-\mu)\}. \tag{5.4}$$

Note that the continuity of the risk measure ρ guarantees that the maximum in problem (5.3) is achieved. We note that by the theorem, $-\partial\rho(-Q(z))$ is a subset of A. With this representation, the optimization problem (5.2) takes the form

$$\min_{z \in Z} \max_{\mu \in A} \{\langle \mu, Q(z)\rangle - \rho^*(-\mu)\} \qquad A = -dom(\rho^*) \tag{5.5}$$

If the problem (5.5) has a non-empty and bounded set of optimal solutions, then we obtain the following duality result [35].

Theorem 11. *Suppose the function $Q : \mathbb{R}^n \to \mathcal{X}$ is convex and the function $\rho : \mathcal{X} \to \mathbb{R}$ is proper, lower semicontinuous and satisfies assumptions $(A_1) - (A_3)$. Suppose further that problem (5.5) has a non empty and bounded set of optimal solutions. Then the* *optimal value of problem* (5.5) *is equal to the optimal value of the problem*

$$\max_{\mu \in A} \min_{z \in Z} \{\langle \mu, Q(z) \rangle - \rho^*(-\mu)\}, \qquad A = -dom(\rho^*). \tag{5.6}$$

Thus the optimal solution of the problem is a saddle point $(\overline{z}, \overline{\mu})$, with

$$\overline{z} \in \arg\min_{z \in Z} \{\langle \mu, Q(z) \rangle - \rho^*(-\mu)\}, \tag{5.7}$$

$$\overline{\mu} \in \arg\max_{\mu \in A} \{\langle \mu, Q(z) \rangle - \rho^*(-\mu)\}. \tag{5.8}$$

Note that equation (5.8) implies that $\overline{\mu}$ is an element of the subdifferential $\partial\rho(-Q(z))$. If, in addition, ρ is positively homogeneous, then ρ^* is the indicator function of the set $U \subset -P$, and the optimization problem takes the form

$$\max_{z \in Z} \min_{\mu \in A} \langle \mu, Q(z) \rangle \tag{5.9}$$

with optimal saddle point solution $(\overline{z}, \overline{\mu})$ and $\overline{\mu} \in \partial\rho(-Q(\overline{z}))$ a probability measure. The optimal $\overline{\mu}$ in the saddle point is the first stage risk-adjusted probability measure. As in the first part of the paper, $\overline{u}$ represents the worst possible outcome in the matrix game (5.9). In the next two sections we theoretically construct $\overline{\mu}$ for the mean-semideviation and mean-deviation from quantile two-stage portfolio problems, in the discrete outcome space.

5.2 The Mean-Semideviation Model

5.2.1 Model Formulation

The mean-semideviation risk model was defined in section 1:

$$\rho_1[X] = -E[X] + \gamma E \max(E[X] - X, 0). \tag{5.10}$$

For the risk aversion constant $\gamma \in [0,1]$, it was shown in [22] that the risk measure $\rho_1(.)$ is coherent. Setting $X = -Q(z^1)$, the risk-averse portfolio problem (5.2) with risk measure (5.10) is given by

$$\min_{z^1 \in Z_1} \langle p^1, Q(z^1) \rangle + \gamma \sum_{i=1}^{m_1} p_i^1 \max(Q_i(z^1) - \langle p^1, Q(z^1) \rangle,\ 0) \tag{5.11}$$

where $Q_i(z^1)$ is the optimal solution to an *ith* second stage problem. Recall that the second stage objective ρ_{2i} in the definition (4.11) of Q_i is a conditional risk mapping . Theoretically, there are many different functions we could use as the second stage objective ρ_{2i}. For this section, we will assume that the second stage objective is also a mean-semideviation function. The second stage problem $Q_i(z^1)$ is then given by

$$\min_{z_i^2 \in Z_{2i}} -\langle p_i^2, X_i^2 \rangle + \gamma \sum_{k=1}^{m_2} p_{ik}^2 \max(\langle p_i^2, X_i^2 \rangle - X_{ik}^2,\ 0) \tag{5.12}$$

where $X_i^2 = \xi_i^2 z_i^2$ represents the second stage end portfolio value. As mentioned earlier, it may be difficult to interpret the meaning of this composite mapping intuitively. However, for illustrative purposes, it can be used to show how first and second stage risk-adjusted measures are constructed.

Using the representation theorem 10, we can write the portfolio problem equivalently as

$$\min_{z^1 \in Z_1} \max_{\mu^1 \in A} \langle Q(z^1), \mu^1 \rangle, \tag{5.13}$$

where $A = -\partial\rho(0)$ is a subset of the set of probability measures on $\{1, 2, ..., m_1\}$. The general representation for $-\partial\rho(-Q(z^1))$ is given from Theorem 10

$$-\partial\rho(-Q(z^1)) = \arg\max_{\mu^1 \in A} \langle \mu^1, Q(z^1) \rangle \tag{5.14}$$

where $A = -\partial\rho(0) \subset P$. To calculate the first stage risk adjusted probability measure, we can solve problem (5.11), obtain the optimal portfolio vector z^1, and solve the following optimization problem

$$\begin{aligned} &\text{Maximize } \langle \mu^1, Q(z^1)\rangle \\ &\text{s.t. } \mu^1 \in -\partial\rho(0). \end{aligned} \tag{5.15}$$

The subdifferential $-\partial\rho(0)$ was given in (2.43) of section 1:

$$-\partial\rho(0) = \{(1-\gamma\langle g,\ 1\rangle)p^1 + \gamma g : |g_i| \leq \frac{1}{2}p_i^1,\ i = 1..m_1\}. \tag{5.16}$$

The *ith* second stage risk-adjusted probability measures μ_{2i} can be calculated as in part 1, by formulating the dual linear problem to problem (5.12) as a matrix game, with optimal μ_{2i} part of the saddle point solution.

We can also obtain closed form expressions for μ^1 and μ_i^2 by directly calculating the subdifferentials of the objective functions in (5.11) and (5.12), respectively, which we now show.

5.2.2 Subdifferentials

Consider the first stage objective function $\rho_1(.)$ in (5.11)

$$\rho(-Q(z^1)) = E_{p^1}[Q(z^1)] + \gamma E_{p^1}[\max(Q(z^1) - E_{p^1}[Q(z^1)],\ 0)]. \tag{5.17}$$

Both the expectation and semideviation functions are proper, convex and finite on $\mathcal{X}$. By the Moreau-Rockafellar theorem, then, the subdifferential of ρ_1 at the point $X_1 = -Q(z^1)$ is given by the sum of the subdifferentials of the two functions at X_1. That is,

$$\partial\rho_1(X_1) = \partial E_{p^1}[Q(z^1)] + \gamma\partial(E_{p^1}[\max(Q(z^1) - E_{p^1}[Q(z^1)],\ 0)) \tag{5.18}$$

The calculation of the subdifferential for the expectation function is straightforward. Since it is differentiable, its subdifferential consists of a single point, the vector p^1. The calculation of the subdifferential for the semideviation function requires more work. To aid in this calculation, we provide an equivalent representation for this function.

Lemma 12.

$$E[\max\{0, X - E[X]\}] = \frac{1}{2}E|X - E[X]|. \tag{5.19}$$

In this form, the semideviation function can be viewed as the p^1-norm of the function $BX = X - E[X]$, where the p^1-norm is defined by

$$\|w\|_{p^1} = \sum_{i=1}^{m} p_i^1 |w_i|. \tag{5.20}$$

The dual norm to $\|w\|_{p^1}$ is

$$\|v\|_* = \max_i |\frac{v_i}{p_i^1}|. \tag{5.21}$$

In fact, the subdifferential of the function $f(BX) = \|BX\|$ was calculated in [34]. We extend that calculation for the semideviation function.

Theorem 13. *The subdifferential of the function $f(BX) = \|BX\|_{p^1}$, where $BX = X - E[X]$ is given by*

$$\partial f(X) = \{h = \frac{1}{2}g - \frac{1}{2}\langle g, p^1\rangle\} \tag{5.22}$$

where g_i are given by

$$g_i = \begin{cases} p_i^1 & \text{if } X_i > E[X] \\ -p_i^1, & \text{if } X_i < E[X] \\ \theta_i \in [-p_i^1, p_i^1], & \text{if } X_i = E[X] \end{cases} \tag{5.23}$$

Proof. The subdifferential for the general norm is [34]

$$\partial\|x\|_\diamondsuit = \{g \in \mathbb{R}^m : \|g\|_* \leq 1, \langle g, x\rangle = \|x\|_\diamondsuit\} \tag{5.24}$$

with $\|.\|_*$ denoting the dual norm. Using (5.20), we have

$$\partial\|x\|_{p^1} = \{g \in \mathbb{R}^m : \|g\|_* \leq 1, \langle g, x\rangle = \langle p^1, x\rangle\}. \tag{5.25}$$

Applying the chain rule for subdifferentials, with $BX = X - E[X]$, we obtain

$$\begin{aligned}\partial\|BX\| &= B^T\{g \in \mathbb{R}^m : \|g\|_* \leq 1, \langle g, BX\rangle = \|BX\|_{p^1}\} \\ &= \{g - \langle g, p^1\rangle : \|g\|_* \leq 1, \langle g, X - EX\rangle = \|X - EX\|_{p^1}\}.\end{aligned} \tag{5.26}$$

Solving equation in (5.26), we obtain (5.23) for g. □

Combining the two subdifferentials, we obtain

$$A = \{u_1 = (1 - \frac{1}{2}\gamma\langle g, 1\rangle)p^1 + \frac{1}{2}\gamma g\} \tag{5.27}$$

with g given by (5.23). Thus, to solve for u_1, we could calculate g using (5.23) and substitute into equation(5.27). In programming models, this approach is faster than solving the subproblem (5.15). To obtain the i^{th} second stage risk adjusted probability measure u_{2i}, we would similarly calculate the subdifferential of the objective function with respect to the variable $X_{2i} = \xi_i^2 z_i^2$

$$-\langle p_i^2, X_{2i}\rangle + \gamma E_{p^2} \max(E[X_i^2] - X_i^2,\ 0). \tag{5.28}$$

The subdifferential of the first term is the vector p_{ik}^2. The calculation of the subdifferential of the second term follows the method given in part 1. We state it here without proof.

Theorem 14. *The subdifferential of the function ρ_{2i} takes the form:*

$$A_2 = \{u_{2i} = (1 - \frac{1}{2}\gamma\langle g_i^2, 1\rangle)p_i^2 + \frac{1}{2}\gamma g_i^2\} \tag{5.29}$$

where g_i^2 is given by

$$g_i = \begin{cases} p_{ik}^2, & \text{if } \; E[X_i^2] > X_{ik}^2 \\ -p_i^2, & \text{if } \; E[X_i^2] < X_{ik}^2 \\ \theta_{2i} \in [-p_i^2, p_{ik}^2], & \text{if } \; E[X_i^2] = X_{ik}^2 \end{cases} \tag{5.30}$$

5.3 The Mean-Weighted Deviation from Quantile Model

5.3.1 Model Formulation

The mean-deviation from quantile risk measure was defined in (2.39) by

$$\rho_1(X) = -E[X] + \gamma E \max\{\frac{1-\alpha}{\alpha}(q_\alpha - X),\ (X - q_\alpha)\}, \tag{5.31}$$

where γ is some risk-aversion constant and q_α is the α-quantile for $\alpha \in [0,1]$. For $\gamma \in [0,1]$, it is known from [35] that ρ_1 is a coherent risk measure. Also recall that

$$E \max\{\frac{1-\alpha}{\alpha}(q_\alpha - X),\ (X - q_\alpha)\} = \min_{\eta \in R} E \max\{\frac{1-\alpha}{\alpha}(\eta - X),\ (X - \eta)\}. \tag{5.32}$$

The optimal η in the second expression of (5.32) is the $\alpha-$quantile of the distribution of X. We formulate now the two-stage discrete portfolio problem. Setting $X = -Q(z^1)$ and $\eta_1 = -\eta$, the first stage problem is

$$\min_{z^1 \in Z_1, \eta_1 \in R} \langle p^1, Q(z^1) \rangle + \gamma \sum_{i=1}^{m_1} p_i^1 \max\{\frac{1-\alpha}{\alpha}(Q_i(z^1) - \eta_1), (\eta_1 - Q_i(z^1))\}, \tag{5.33}$$

where the $Q_i(z^1)$ is taken as the optimal value of the following second stage problem,

$$\min_{z_i^2 \in Z_{2i}, \eta_{2i} \in \mathbb{R}} -\sum_{k=1}^{m_2} p_{ik}^2 X_{ik}^2 + \gamma \sum_{k=1}^{m_2} p_{ik}^2 \max\{(\frac{1-\alpha_2}{\alpha}(\eta_{2i} - X_i^2)\ , (X_i^2 - \eta_{2i})\}. \tag{5.34}$$

Here η_{2i} is the $\alpha-$ quantile for the second stage random variable X_i^2. Using the representation (2.11) from the previous section, problem (5.33) can be written equivalently as

$$\min_{z^1 \in Z_1} \sup_{\mu^1 \in A} \langle Q(z^1), \mu^1 \rangle \tag{5.35}$$

where $A = -\partial\rho(0)$ is a subset of the set of probability measures on $\{1, .., m_1\}$. A general representation for $-\partial\rho(-Q(z^1))$ is given from Theorem 10 :

$$\arg \max_{\mu^1 \in -\partial\rho(0)} \langle \mu^1, Q(z^1) \rangle. \tag{5.36}$$

Thus we can calculate the first stage risk-adjusted probability measures by solving problem (5.33), obtaining the optimal portfolio vector z^1, and solving the following optimization problem

$$\begin{aligned} &\text{Maximize } \langle \mu^1, Q(z^1) \rangle \\ &\qquad\quad \text{s.t. } \mu^1 \in -\partial\rho(0). \end{aligned}$$

Using previous results, the subdifferential $\partial\rho(0)$ is

$$-\partial\rho(0) = \{(1-\gamma)p^1 + \gamma g : 0 \leq g_i \leq \frac{p_i^1}{\alpha},\ i = 1..m_1,\ \sum_{i=1}^{m_1} g_i = 1\}. \tag{5.37}$$

We can also obtain a closed form expression for μ_1 by directly calculating the subdifferential of the objective function (5.33), which we now show.

5.3.2 Subdifferentials

Consider again the objective function ρ_1 given in (5.33).

$$\rho_1(Q(z^1)) = E[Q(z^1)] + \gamma E[\max\{\frac{1-\alpha}{\alpha}(Q_i(z^1) - \eta_1), (\eta_1 - Q_i(z^1))\}] \tag{5.38}$$

Both the expectation and mean-deviation from quantile functions in (5.38) are proper, convex and finite on $\mathcal{X}$. By the Moreau-Rockafellar theorem, then, the subdifferential of ρ_1 at the point $X^1 = Q(z^1)$ is given by the sum of the subdifferentials of the two functions at X^1. That is,

$$\partial\rho_1(X^1) = \partial E_{p^1}[X^1] + \gamma\partial(E\max\{\frac{1-\alpha}{\alpha}(\eta_1 - X^1),\ (X^1 - \eta_1)\}). \tag{5.39}$$

The subdifferential is given below:

Theorem 15. *The subdifferential of the mean-deviation from quantile function is given by*

$$\partial\rho_1(X^1) = (1-\gamma)p^1 + \gamma g_1 \tag{5.40}$$

where

$$g_{1i} = \begin{cases} (\frac{p_i^1}{\alpha}), & \textit{if} \ \ X_i^1 > \eta_1 \\ 0, & \textit{if} \ \ X_i^1 < \eta_1 \\ \theta_i \in [0, \frac{p_i^1}{\alpha}], & \textit{if} \ \ X_i^1 = \eta_1 \end{cases} \tag{5.41}$$

and with $\sum_{i=1}^{m_1} g_{1i} = 1$.

Proof. Simplifying the objective function ρ_1, we obtain

$$\rho(X) = (1-\gamma)E[X^1] + \min_{\eta_1 \in \mathbb{R}} \sum_{i=1}^{m_1} p_i^1 \max\{\frac{1}{\alpha}(X_i^1 - \eta_1), 0\}. \tag{5.42}$$

There are three main cases to consider. If $X_{1i} > \eta_1$, then the subdifferential can be determined by differentiating the term

$$\frac{p_i^1}{\alpha} X_i^1 \tag{5.43}$$

We differentiate in a similar way for the cases $X_i^1 < \eta_1$. Note that in the discrete case, we can apply the following equation to make u_1 a probability measure. Letting the case where $X_i^1 = \eta_1$ be denoted by k, we obtain

$$g_{k1} = 1 - \sum_{i \neq k} g_{i1}. \tag{5.44}$$

Thus, the first stage risk-adjusted probability measure would take the form

$$\mu_1 = (1 - \gamma)p^1 + \gamma g_1 \tag{5.45}$$

with g_{1i} given by (5.41). $\square$

Chapter 6

Benders' Decomposition

6.1 Review of Benders' Decomposition

In the previous sections, we formulated the discrete two-stage risk-averse portfolio problem

$$\min_{z^1 \in Z_1} c^T z^1 + \rho_1(-Q(z^1)). \tag{6.1}$$

The main representation theorem and optimality conditions were presented for problem (6.1) and we derived first and second stage risk-adjusted probability measures. Our focus now shifts to solution methods for problem (6.1).

In order to accomplish this, we recall solution methods used for the discrete linear two-stage stochastic programming problem given in [39]

$$\begin{aligned} \text{Minimize } & c^T z^1 + E[Q(z^1)] \\ \text{s.t. } & Az^1 = b, \quad z^1 \geq 0 \end{aligned} \tag{6.2}$$

where $Q_i(z^1)$ is the optimal value of the ith second stage problem

$$\begin{aligned} \text{Minimize } & q_i^T z_i^2 \\ \text{s.t. } & T_i z^1 + W_i z^2 = h_i, \quad z_i^2 \geq 0 \end{aligned} \tag{6.3}$$

There are two main classes of solution methods described for problem (6.2) - (6.3), primal decomposition methods and dual methods [40]. As described in [39], the primal decomposition methods solve many subproblems of the form (6.3) to construct approximations for $Q_i(z^1)$ and for the expectation of Q. These approximations are used in a master problem which generates approximations for the first stage solution z^1.

In this section, we focus on the primal decomposition methods. One of the most important methods in this class is the cutting plane method, or Benders' Decompostion, which we describe in more detail below. The objective for this section will be to extend the cutting plane method for problem (6.2) - (6.3) to solving the two-stage risk averse portfolio problem (6.1). We will also review and extend another composite multicut version of this method.

The idea of Benders' Decomposition method is to construct a sequence of approximations $\{z_k^1\}$ to the solution of problem (6.2). At each iteration k, the method attempts to solve all second stage subproblems (6.3) to generate models for $Q_i(z^1)$ and $E[Q(z^1)]$. If successful, the model for $E[Q(z^1)]$ is added to a master problem and a new approximation z_{k+1}^1 is generated. If an infeasibile second stage problem is encountered, then the method attempts to cut off the approximation from future consideration. We describe now in detail these two cases.

Suppose first that all second stage problems (6.3) are feasible for approximation z_k^1. To construct a model for $Q_i(z^1)$, the method generates a hyperplane which bounds the function from below. From convex analysis, such a hyperplane takes the form

$$Q_i(z^1) \geq Q_i(z_k^1) + \langle \Psi_i^k, z^1 - z_k^1 \rangle \tag{6.4}$$

where $\Psi_i^k \in \partial Q_i(z_k^1)$ is a subgradient of Q_i at the point z_k^1. The expression for the subdifferential [40] $\partial Q_i(z_k^1)$ is given by

$$\partial Q_i(z_k^1) = -(T^i)^T D_i^k(z_k^1) \tag{6.5}$$

where

$$D_i^k(z_k^1) := \arg \max_{W_i^T \sigma_i \leq q_i} \sigma_i^T (h_i - T_i z_k^1) \tag{6.6}$$

is the set of optimal solutions to the dual problem of (6.3). Letting σ_i^k be one of the optimal solutions in the set (6.6), we have

$$Q_i(z^1) \geq Q_i(z^1_k) - \langle (T^i)^T \sigma_i^k, z^1 - z^1_k \rangle. \tag{6.7}$$

The hyperplane in (6.7) is called an optimality cut for $Q_i(.)$. The optimality cut for $Q(z^1)$ can be determined by taking the following sum

$$Q(z^1) = \sum_{i=1}^{m_1} p_i^1 Q_i(z^1) \geq \sum_{i=1}^{m_1} p_i^1 (Q_i(z^1_k) - \langle (T^i)^T \sigma_i^k, z^1 - z^1_k \rangle. \tag{6.8}$$

Suppose now at iteration k the approximation $\{z^1_k\}$ is infeasible to the ith second stage problem (6.3). We show now how to construct a feasibility cut at $\{z^1_k\}$. Recall that to check feasibility of a linear program, the following phase 1 problem is constructed

$$\begin{aligned} &\text{Minimize } \|x\| \\ &\qquad \text{s.t. } W_i z_i^2 + x = h_i - T^i z^1 \\ &\qquad\qquad z_i^2 \geq 0. \end{aligned} \tag{6.9}$$

Here, $x = (x_1, .., x_m)$ is a vector of artificial variables, and the term $\|x\|$ denotes the norm of x on the space R^m. The l_1 norm $\|x\|_1 = |x_1| + ... + |x_m|$ or the l_∞ norm $\|x\|_\infty = \max(|x_1|, .., |x_m|)$ will normally be used. Both norms are polyhedral functions, in the sense that they can be represented as the maximum of a finite number of linear functions and thus the problem above has a linear programming representation. This proves to be an important fact in the convergence proof for Benders Decomposition.

Let $U^i(x)$ denote the optimal value of the primal problem 6.9. If $U^i(x) = 0$, then the second stage problem would be feasible. Thus, for infeasible z^1_k, we have that $U^i(z^1_k) > 0$. The dual problem to 6.9 is

$$\max_{\sigma_i^k} \min_{z_i^2 \in Z_{2i}, x} \{\|x\| + (h^i - T^i z^1 - W^i z^2 - x)^T \sigma_i^k\}. \tag{6.10}$$

Calculating the minimum of the inside term, the problem becomes

$$\text{Maximize } (h_i - T^i z^1)^T \sigma_i^k$$
$$\text{s.t. } W_i^T \sigma_i^k \leq 0$$
$$\|\sigma_i^k\|_* \leq 1,$$

where $\|.\|_*$ is the dual norm to $\|.\|$. The primal and dual problems will have the same optimal value by duality theory of linear programming. If the second stage problem is infeasible, then the dual objective value will be positive. Thus the feasibility cut will take the form

$$(h_i - T^i z^1)^T \sigma_i^k \leq 0 \tag{6.11}$$

Similar considerations can be carried out for a general polyhedral set Z_{2i}.

In both cases, we add the feasibility cut (or optimality cut) to a master problem consisting of previous feasibility and optimality cuts, and the original constraints of (6.2). The solution of the master problem at iteration k yields a new approximation, $\{z_{k+1}^1\}$, and the process repeats until an optimality condition is satisfied.

Having determined the optimality and feasiblity cuts, the master problem can be formulated. Let k denote the iteration number, and let J_{opt} denote the set of iteration numbers corresponding to the construction of an optimality cut. Such an iteration is called an outer iteration. Let $J_{feas}(i)$ denote the set of iterations where a feasibility cut is constructed for scenario i. Such an iteration is called an inner iteration. The master problem takes the form

$$\begin{aligned}
&\text{Minimize } c^T z^1 + v \\
&\quad \text{s.t. } v \geq \sum_{i=1}^{m_1} p_i^1 (Q_i(z_k^1) - \langle (T^i)^T \sigma_i^k, z^1 - z_k^1 \rangle \quad k \in J_{\text{opt}} \\
&\quad \text{s.t. } (h_i - T^i z^1)^T \sigma_i^k \leq 0 \ \ \forall i = 1..m_1, \quad k \in J_{\text{feas}}(i) \\
&\quad \text{s.t. } Az^1 = b, \ z^1 \geq 0
\end{aligned}$$

The solution to the master problem, z^{k+1} and v^{k+1} is used as the next approximation to the two-stage problem, and to construct optimality or feasibility cuts for Q. Optimality occurs when the new cut does not cut the current solution off. We summarize the method below.

STEP 0: Initialize $k = 0$

STEP 1: Set $k = k + 1$, solve the master problem. Let (z^k, v^k) be optimal solution.

STEP 2: Using $z = z^k$, check all second stage primal problems for feasibility. If yes, go to step 3. Otherwise, locate the first problem that does not have feasible solution. Generate a feasibility cut, add feasibility cut to the master problem, and return to step 1. Repeat until all secondnd stage primal problems are feasible.

STEP 3: Check for the optimality condition. If it is not satisfied, obtain optimality cut using subdifferentials. Add optimality cut to master problem. Return to step 1.

6.2 Extension of Benders' Decomposition

With the cutting plane method for the linear two-stage problem (6.2) - (6.3) developed, we focus on extending this method to the two stage risk-averse problem (6.1). For simplicity, the sets Z_1 and Z_{2i} are taken from the previous linear problem. That is, $Z_1 = \{z^1 : Az^1 = b, \;\; z^1 \geq 0\}$ and $Z_{2i} = \{z_i^2 : W^i z_i^2 = h^i - T^i z^1, z_i^2 \geq 0\}$.

In order to extend the cutting plane method, we need to develop optimality cuts for the objective function in (6.1) and feasibility cuts for the sets Z_{2i}. As the sets Z_{2i} are identical to those in the linear case, the process for constructing the feasibility cuts is the same. We focus now on constructing optimality cuts.

Suppose we denote by $\phi(z^1)$ the objective function in (6.1). As in the linear case, we want to construct a model for $\phi(z^1)$ using the information at z_k^1. That is, to construct a hyperplane bounding $\phi(z_k^1)$ from below. Using convex analysis as in the previous case, the optimality cut will be of the form

$$\phi(z^1) \geq \phi(z_k^1) + \langle \Psi^k, z^1 - z_k^1 \rangle \tag{6.12}$$

where $\Psi^k \in \partial\phi(z_k^1)$ is an element of the subdifferential for ϕ at z_k^1.

Thus, it is neccessary to determine the form of the subdifferential $\partial\phi(z^1)$ for an arbitrary coherent risk measure $\phi(z^1)$. In order to illustrate some of the technical issues that arise, we compare the function $\phi(z^1)$ with the expectation function

$$E[Q(z^1)] = \sum_{i=1}^{m_1} p_i^1 Q_i(z^1) \tag{6.13}$$

In equation (6.13), the function $E[Q(z^1)]$ is written as the weighted sum of the objective functions $Q_i(z^1)$. As was illustrated in the previous section, the calculation of the subdifferential at the point z^1 in this case is straightforward : we simply add the weighted sum of the subdifferentials of the individual functions $Q_i(.)$

$$\partial E[Q(z^1)] = \sum_{i=1}^{m_1} p_i^1 \partial Q_i(z^1) \tag{6.14}$$

Moreover, the second stage problems are themselves linear programming problems, so we can determine a closed form for the subdifferential $\partial Q_i(z^1)$ by calculating the optimal solution vector of the dual problem. In the general case, however, the function $\phi(z^1)$ may have a much more complicated structure. For example, if we let ρ_1 be the semideviation function, then,

$$\phi(z^1) = \rho_1(-Q(z^1)) = E[Q(z^1)] + \sum_{i=1}^{m_1} p_i^1 \max(Q_i(z^1) - E[Q(z^1)], 0) \tag{6.15}$$

Calculating the subdifferential of $\rho_1(.)$ with respect to $X = Q(z^1)$ required some effort. We had to rewrite the function ρ_1 as

$$\frac{1}{2}\sum_{i=1}^{m_1} p_i^1 |Q_i(z^1) - E[Q(z^1)]| = \frac{1}{2}\|BQ(z^1)\|_{p^1} \tag{6.16}$$

Even in this form, we had to use chain rules of subdifferentials. Moreover, the second stage functions $Q_i(z^1)$ are not linear either, so it is not at all clear how to obtain a closed form as in (6.13). The function $\partial Q_i(z^1)$ has its own computed form. In order to obtain a closed form for the expression $\partial\phi(z^1)$ we will require more sophisticated methods.

Before proceeding with this calculation, we provide some background notation. Recall that the function $Q : \mathbb{R}^n \to \mathcal{X}$ is a random function with m_1 realizations Q_i corresponding to outcome i. Its subdifferential will also be a random set of vectors, with realizations $\Psi_i \in \partial Q_i(z^1)$, corresponding to outcome i. We denote by s_i a subgradient belonging to the ith subdifferential $\partial Q_i(z^1)$. Note that s_i is a vector in $\mathbb{R}^n$. We also introduce the notation $E_p(s)$ to represent the sum $\sum_{i=1}^{m_1} p_i s_i$.

The theorem for calculating the subdifferential is given below.

Theorem 16. *Let $\mathcal{X} := L_p(\Omega, F, P)$ and $Q : \mathbb{R}^n \to \mathcal{X}$ be a convex mapping. Suppose that the mapping ρ_1 satisfies conditions (A_1) and (A_2) and is finite-valued and continuous at $X = -Q(z^1)$. Then $\phi = \rho_1(-Q(z^1))$ is subdifferentiable at z^1 and*

$$\partial\phi(z^1) = \{E_\mu(s)|\mu \in -\partial\rho_1(-Q(z^1)),\ s \in \partial Q(z^1)\ \}. \tag{6.17}$$

We make a few notes regarding equation (6.17). It represents the set of vectors $g \in \mathbb{R}^n$ representable as the following sum

$$g = \sum_{i=1}^{m_1} \mu_i s_i \tag{6.18}$$

where s_i is a vector in R^n and μ_i is a real number. Recall that a subgradient $\mu \in -\partial\rho_1(-Q(z^1))$ for ρ_1 coherent can be calculated from the maximization problem (5.14), and is a probability measure.

Thus to calculate a particular subgradient Ψ^k at iteration k, we first solve the maximization problem (5.14) to determine an optimal probability measure μ^k. We then obtain the subdifferential $\partial Q_i(z^1)$ for each i to obtain a vector s_i, and take their sum given in (6.18). The optimality cut takes the form

$$\phi(z^1) \geq \phi(z_k^1) + \langle E_{\mu^k}(s), z^1 - z_k^1\rangle, \tag{6.19}$$

With the feasibility and optimality cuts constructed, the master problem at iteration k is given by

$$\begin{aligned}
&\text{Minimize } v \\
&\qquad \text{s.t. } v \geq \phi(z_k^1) + \langle E_{\mu^k}(s), z^1 - z_k^1\rangle, k \in J_{opt} \\
&\qquad \text{s.t. } (h_i - T_i z^1)^T \sigma_i^k \leq 0 \quad \forall i = 1..m_1, \quad k \in J_{feas}(i) \\
&\qquad \text{s.t. } Az^1 = b,\ z^1 \geq 0
\end{aligned} \tag{6.20}$$

Optimality occurs when the new cut does not cut the current solution off. In summary, the Extended Bender's decomposition method adapts the standard method to solving the larger class of two-stage stochastic programming problems with coherent risk objectives. In particular, we can apply this method to coherent mean-risk models of the form

$$\min_{z^1 \in Z_1} E[Q(z^1)] + \gamma r[-Q(z^1)] \tag{6.21}$$

where r is a coherent risk functional. In what follows, we formulate explicitly the feasibility and optimality cuts for the discrete two-stage mean-semideviation and mean-deviation from quantile models.

6.3 Mean-Risk Models

We construct the feasibility and optimality cuts for the discrete version of two-stage mean-risk portfolio problems with risk functional defined to be the semideviation. Consider the *ith* second stage problem

$$\min_{z_i^2 \in Z_{2i}} \{\sum_{k=1}^{m_2} p_{ik}^2 X_{ik}^2 + \gamma \sum_{k=1}^{m_2} p_{ik}^2 \max(\sum_{l=1}^{m_2} p_{il}^2 X_{il}^2 - X_{ik}^2,\ 0)\} \tag{6.22}$$

with $X_i^2 = z_i^2 \xi_i^2$. Recall from 4.6that the second stage portfolio problem with domain Z_{2i} has the same optimal objective value as the portfolio problem with $\overline{Z}_{2i}$ given by

$$\begin{aligned} \overline{Z}_{2i} = \{(z_i^2, u_i, v_i) : \sum_{j=1}^{n} \xi_j^1 z_j^1 - \kappa \sum_{j=1}^{n} (u_{ji} + v_{ji}) \geq \sum_{j=1}^{n} z_{ji}^2 \\ u_{ji} - v_{ji} = \xi_j^1 z_j^1 - z_{ji}^2, \quad z_i^2, u_i, v_i \geq 0 \quad \kappa \in [0,1]\}. \end{aligned}$$

Observe that the set $\overline{Z}_{2i}$ is non-empty for every first stage portfolio allocation z^1. Indeed, the option always exists not to rebalance the portfolio, in which case both u_i and v_i are the zero vectors. Thus, all second stage problems will be feasible for every approximation $\{z_k^1\}$, and there is no need to construct feasibility cuts.

We focus now on constructing the optimality cuts for this problem

$$v \geq \phi(z_k^1) + \langle \partial\phi(z_k^1), z^1 - z_k^1 \rangle. \tag{6.23}$$

From the previous section, to determine the subdifferential $\partial\phi(z_k^1)$, we need to determine the subdifferentials $\partial Q_i(z_k^1)$ of the second-stage problems, and the first-stage risk-adjusted measures μ_{ik}^1. The first-stage risk adjusted measures were constructed in the previous section. To construct the subdifferentials, consider the second-stage mean-semideviation portfolio problem in scenario i :

$$\begin{aligned}
\text{Minimize } & -\sum_{k=1}^{m_2} p_{ik} \sum_{j=1}^{n} \xi_{jik}^2 z_{ji}^2 + \gamma \sum_{k} p_{ik} \max((\sum_{l} p_{il} \sum_{j=1}^{n} \xi_{jil}^2 z_{ji}^2 - \sum_{j=1}^{n} \xi_{jik}^2 z_{ji}^2), 0) \\
\text{s.t. } & \sum_{j=1}^{n} z_{ji}^2 + \kappa \sum_{j=1}^{n} (u_{ji} + v_{ji}) = \sum_{j=1}^{n} z_j^1 \xi_{ji}^1 \\
\text{s.t. } & z_{ji}^2 - u_{ji} + v_{ji} = z_j^1 \xi_{ji}^1, \qquad \forall j = 1..n \\
& z_{ji}^2 \geq 0 \;\; v_{ji} \geq 0, \;\; u_{ji} \geq 0.
\end{aligned}$$

Using Lagrangian duality, we can determine the expression for the dual problem

$$\begin{aligned}
\text{Maximize } & \sum_{j=1}^{n} (\lambda_{ji} - \lambda_{0i})(z_j^1 \xi_{ji}) \\
\text{s.t. } & [\ (\sum_{k=1}^{m_2} (t_i p_{ik} - \lambda_{3ik}) \xi_{jik} + (\lambda_{0i} - \lambda_{ji})\ ] \geq 0, \qquad \forall j \\
& \gamma p_{ik} - \lambda_{3ik} \geq 0 \\
& \kappa \lambda_{0i} + \lambda_{ji} \geq 0 \forall k = 1..m_2 \\
& \kappa \lambda_{0i} - \lambda_{ji} \geq 0 \forall j = 1..n \\
\text{s.t. } & \lambda_{3ik} \geq 0 \;\; \forall k = 1..m_2
\end{aligned}$$

where $t_i = \sum_{l=1}^{m_2} \lambda_{3il}$. The subdifferential of $Q_i(z^1)$ can be calculated as the derivative of the objective function,

$$\partial Q_i = (-\lambda_0 + \lambda)\xi_i \tag{6.24}$$

Since the mean-deviation from quantile portfolio problem has the same feasibility set $\overline{Z}_{2i}$, the objective function of its dual will have the same form as the semideviation model, and thus its subdifferential will be of the same form. However, the Lagrange multipliers $(\lambda_0, \lambda, \lambda_3)$ will take different values.

Thus the optimality cuts take the form

$$\phi(x) \geq \phi(z_k^1) + \sum_{i=1}^{m} u_{ki}^1 \langle (-\lambda_0 + \lambda)_i \xi_i, z^1 - z_k^1 \rangle \tag{6.25}$$

We prove the convergence of this modified method. Providing it can be proved that ρ_1 and ρ_2 in the above cases are polyhedral, the proof will follow a similar form to that given in [16].

Theorem 17. *Assume the set Z_1 is bounded and $\rho_1(Q(z^1)) < \infty$. Moreover, let all cuts constructed in steps 1 and 2 be basic objective and feasibility cuts. Then after finitely many iterations, the extended Bender's decomposition algorithm finds an optimal solution to the two-stage risk averse portfolio problem.*

Proof. Suppose that at iteration k, a basic feasibility cut has been constructed for z_k^1. The feasibility cut removes z_k^1 from future solutions to the master problem. Since the number of basic feasible solutions to (2.8) is finite, for each *ith* second stage problem, the number of basic feasibility cuts from the *ith* problem is finite. Hence, the total number of feasibility cuts is finite.

The second stage objective function Q_i is polyhedral. Thus, the ith second stage problem can be written as a linear programming problem. The optimal solution to the dual problem provides the lagrange multiplier λ_{ik} for constructing the optimality cut in (6.25). If we assume that λ_{ik} is basic, then there are finitely many basic feasible solutions, hence only finitely many optimality cuts for each Q_i. Since the optimality cut for ρ_1 is the expectation of the optimality cuts for Q_i with respect to risk adjusted probability measures, there are finitely many optimality cuts for ρ_1. $\square$

6.4 Multi-Cut Benders' Decomposition

In this section, we describe the multi-cut version of Benders' Decomposition on the discrete linear two-stage problem (6.2) - (6.3), and extend it to the general discrete two-stage problem (6.1) with a coherent risk measure.

The multi-cut version of Benders' Decomposition differs from the standard version in the way it constructs optimality cuts. Recall in the standard version, that the hyperplane bounding $Q_i(z^1_k)$ from below was constructed for each outcome i by calculating the subdifferentials $\partial Q_i(z^1_k)$. The optimality cut at iteration k was then constructed by taking the average of these hyperplanes

$$v \geq \sum_{i=1}^{m_1} p^1_i (Q_i(z^1_k) + \langle \psi^k_i, z^1 - z^1_k \rangle, \quad \psi^k_i \in \partial Q_i(z^1_k) i = 1..m_1, \quad k \in J_{opt} \tag{6.26}$$

In the multi-cut version, these hyperplanes are not averaged, but taken as m_1 separate optimality cuts corresponding to the m_1 second stage problems

$$v^i \geq Q_i(z^1_k) + \langle \Psi^k_i, z^1 - z^1_k \rangle, \quad i = 1..m_1, \quad k \in J_{opt}, \ \Psi^k_i \in \partial Q_i(z^1_k) \tag{6.27}$$

The master problem thus takes the form

$$\begin{aligned} &\text{Minimize } c^T z^1 + \sum_{i=1}^{m_1} p^1_i v^i \\ &\qquad \text{s.t. } v^i \geq Q_i(z^1_k) + \langle \partial \Psi^k_i, z^1 - z^1_k \rangle, \quad i = 1..m_1, \quad k \in J_{opt} \\ &\qquad \text{s.t. } (h_i - T_i z^1)^T \sigma^k_i \leq 0 \ \ \forall i = 1..m_1, \quad k \in J_{feas}(i) \\ &\qquad \text{s.t. } Az^1 = b, \ z^1 \geq 0 \end{aligned} \tag{6.28}$$

6.5 Multicut Risk Decomposition

In this section, we extend the multi-cut Benders' Decomposition method to two-stage stochastic programming problems with coherent risk measures. The two-stage discrete stochastic programming problem with coherent risk objective is given by

$$\min_{z^1 \in Z_1} c^T z^1 + \rho_1(-Q(z^1)), \tag{6.29}$$

where ρ_1 is a coherent risk functional and $Q(z^1)$ is the optimal objective value of the second stage problem. Recall that the 2-stage problem can be written equivalently as

$$\min_{z^1 \in Z_1} \sup_{\mu \in A} \langle \mu, Q(z^1) \rangle \tag{6.30}$$

were $A \subset P$. Letting $v = \sup_{\mu \in A} \langle \mu, Q(z^1) \rangle$, problem (6.30) can be written as

$$\begin{aligned} &\text{Minimize } v \\ &\qquad \text{s.t. } v \geq \langle \mu, Q(z^1) \rangle, \forall \mu \in A \\ &\qquad \text{s.t. } z^1 \in Z_1, \end{aligned}$$

To extend the multi-cut plane method, the individual cutting planes for $Q_i(z^1)$ must be determined. The general form for coherent $Q_i(z^1)$ is given in (6.19). The master problem for the multi-cut method is given by

$$\begin{aligned} &\text{Minimize } v \\ &\qquad \text{s.t. } v \geq \sum_{i=1}^{m} \mu_{i1}^k Q_i(z^1), \quad \forall \mu_1^k \in A^k, \\ &\qquad \text{s.t. } Q_i(z^1) \geq Q_i(z_k^1) + \langle \Psi_i^k, z^1 - z_k^1 \rangle \, \forall i = 1..m, \; k \in J_{opt} \\ &\qquad \text{s.t. } (h_i - T_i z^1)^T \sigma_i^k \leq 0, \quad \forall i = 1..m_1, \quad k \in J_{feas}(i) \\ &\qquad \text{s.t. } z^1 \in Z_1 \end{aligned} \tag{6.31}$$

with $A^k \subset A$.

Comparing the linear master problem (6.28) to the risk-averse master problem (6.31), we note that in the first constraint, the original probability measure p^1 is replaced by the risk-adjusted probability measure μ^1. We can easily extend this in particular to the mean-risk models with risk functionals defined as semideviation and deviation from quantile.

Theorem 18. *Assume the set Z_1 is bounded and $\rho_1(Q(z^1)) < \infty$. Moreover, let all cuts constructed in steps 1 and 2 be basic objective and feasibility cuts. Then after finitely many iterations, the extended Multi-Cut Bender's Decomposition algorithm finds an optimal solution to problem.*

Proof. The proof for convergence of the extended multi-cut Benders' Decomposition is similar to that for the extended standard method. Since the number of basic feasibility cuts for each *ith* second stage problem is finite, and there are finite number of second stage problems, the total number of basic feasibility cuts is finite.

The second stage objective function Q_i is polyhedral. Thus, the ith second stage problem can be written as a linear programming problem. The optimal solution to the dual problem provides the Lagrange multiplier $\lambda^{ik'}$ for constructing the optimality cut in 6.26. If we assume that $\lambda^{ik'}$ is basic, then there are finitely many basic feasible solutions, hence only finitely many optimality cuts for each Q_i. Since there are finitely many Q_i, the total number of basic optimality cuts is finite. □

6.6 Linear Model

In this section, we formulate the the two stage mean-risk portfolio problem equivalently as one large linear programming problem, for the risk functionals semideviation and mean weighted deviation from quantile. The linear problem is solved using the simplex method in the numerical experiments and is compared to the extended Benders' and Multi-Cut Benders methods.

6.6.1 Semideviation

Recall the mean-semideviation risk averse portfolio problem:

$$\min_{z^1 \in Z_1} \langle p^1, Q(z^1) \rangle + \gamma \sum_{i=1}^{m_1} p_i^1 \max(Q_i(z^1) - \langle p, Q(z^1) \rangle,\ 0) \tag{6.32}$$

with $Q_i(z^1)$ the optimal objective value of the ith second stage problem

$$\min_{z_i^2 \in Z_{2i}} -\langle p_i^2, X_{2i} \rangle + \gamma \sum_{k=1}^{m_2} p_{ik}^2 \max(\langle p_i^2, X_{2i} \rangle - X_{2ik},\ 0) \tag{6.33}$$

Here ($X_{2i} = \xi_i^2 z_i^2$) represents the second stage end portfolio value. We convert problem (6.32) - (6.33) to a linear programming problem. Consider first the objective function in (6.32). Denoting the term $\max(Q_i(z^1) - \langle p, Q(z^1) \rangle,\ 0)$ by s_i^1, we can rewrite problem (6.32) as

$$\begin{aligned} \text{Minimize } & \langle p^1, Q(z^1) \rangle + \gamma \sum_{i=1}^{m_1} p_i^1 s_i^1 \\ \text{s.t } \ & s_i^1 \geq Q_i(z^1) - \langle p^1, Q(z^1) \rangle,\ i = 1..m_1 \\ & s^1 \geq 0,\ z^1 \in Z_1 \end{aligned} \tag{6.34}$$

Consider now the second stage problem (6.33).

Denoting the term $\max(\langle p_i^2, X_{2i}\rangle - X_{2ik},\ 0)$ by s_{ik}^2, we can rewrite (6.33) as

$$\begin{aligned} \text{Minimize } & -\langle p_i^2, X_{2i}\rangle + \gamma \sum_{k=1}^{m_2} p_{ik}^2 s_{ik}^2 \\ \text{s.t } & s_{ik}^2 \geq \langle p_i^2, X_{2i}\rangle - X_{2ik} \ i = 1..m_1,\ k = 1..m_2 \\ & s_i^2 \geq 0, \quad X_{2i} = \xi_i^2 z_i^2, \quad z_i^2 \in Z_{2i},\ i = 1..m_1. \end{aligned} \tag{6.35}$$

By representing $Q_i(z^1)$ by variables q_i and adding the inequality

$$q_i = -\langle p_i^2, X_{2i}\rangle + \gamma \sum_{k=1}^{m_2} p_{ik}^2 s_{ik}^2 \tag{6.36}$$

we can combine (6.34) and (6.35) into one large linear programming problem

$$\begin{aligned} \text{Minimize } & \langle p, q\rangle + \gamma \sum_{i=1}^{m_1} s_i^1 \\ \text{s.t } & s_i^1 \geq q_i - \langle p, q\rangle,\ i = 1..m_1 \\ & s_i^1 \geq 0, i = 1..m_1,\ z^1 \in Z_1 \\ & q_i = -\langle p_i^2, X_{2i}\rangle + \gamma \sum_{k=1}^{m_2} p_{ik}^2 s_{ik}^2 \ i = 1..m_1,\ k = 1..m_2 \\ & s_{ik}^2 \geq \langle p_i^2, X_{2i}\rangle - X_{2ik} \ i = 1..m_1,\ k = 1..m_2 \\ & s_i^2 \geq 0, \quad X_{2i} = \xi_i^2 z_i^2, \quad z_i^2 \in Z_{2i},\ i = 1..m_1. \end{aligned} \tag{6.37}$$

6.6.2 Mean Weighted Deviation from Quantile

Recall the mean-deviation from quantile portfolio problem

$$\min_{z^1 \in Z_1, \eta_1 \in \mathbb{R}} \langle p, Q(z^1)\rangle + \gamma \sum_{i=1}^{m_1} p_i^1 \max\{\frac{1-\alpha}{\alpha}(Q_i(z^1) - \eta_1), (\eta_1 - Q_i(z^1))\}, \tag{6.38}$$

where the $Q_i(z^1)$ is taken as the optimal value of the following second stage problem

$$\min_{z_i^2 \in Z_{2i}, \eta_{2i} \in \mathbb{R}} -\sum_{k=1}^{m_2} p_{ik}^2 X_{2ik} + \gamma \sum_{k=1}^{m_2} p_{ik}^2 \max\{(\frac{1-\alpha_2}{\alpha}(\eta_{2i} - X_{2i}) \ , (X_{2i} - \eta_{2i})\} \tag{6.39}$$

Here η_2 is the $\alpha-$quantile for the second stage random variable X_2. We convert the problem (6.38) - (6.39) to a linear programming problem. Consider first the objective function in (6.38). Denoting by u_{1i} and v_{1i} the shortfalls $(Q_i(z^1)-\eta_1)$ and $(\eta_1-Q_i(z^1))$, respectively, we can rewrite (6.38) as

$$\begin{aligned} &\text{Minimize } \langle p, Q(z^1)\rangle + \gamma\sum_{i=1}^{m_1} p_i^1(\frac{1-\alpha}{\alpha}u_{1i}+v_{1i}) \\ &\text{s.t} \;\; u_{1i}-v_{1i} = Q_i(z^1)-\eta_1,\; i=1..m_1 \\ &z^1\in Z_1,\; \eta_1\in\mathbb{R}, u_1\geq 0,\; v_1\geq 0. \end{aligned} \tag{6.40}$$

Consider now the second stage problem (6.39). Denoting the by u_{2ik} and v_{2ik} the shortfalls $(\eta_{2i}-X_{2i})$ and $(X_{2i}-\eta_{2i})$, respectively, we can write (6.39) as :

$$\begin{aligned} &\text{Minimize } -\sum_{k=1}^{m_2} p_{ik}^2 X_{2ik} + \gamma\sum_{k=1}^{m_2} p_{ik}^2(\frac{1-\alpha_2}{\alpha}u_{2ik}+v_{2ik}), \quad i=1..m_1,\; k=1..m_2 \\ &\text{s.t} \;\; u_{2ik}-v_{2ik} = \eta_{2i}-X_{2i},\; i=1..m_1,\; k=1..m_2 \\ &u_{2i}\geq 0,\; v_{2i}\geq 0,\quad z_i^2\in Z_{2i},\; \eta_{2i}\in\mathbb{R},\; i=1..m_1. \end{aligned} \tag{6.41}$$

By By representing $Q_i(z^1)$ by variables q_i and adding the equality

$$q_i = -\sum_{k=1}^{m_2} p_{ik}^2 X_{2ik} + \gamma\sum_{k=1}^{m_2} p_{ik}^2(\frac{1-\alpha_2}{\alpha}u_{2ik}+v_{2ik}) \tag{6.42}$$

we can combine (6.40) and (6.41) into one large linear programming problem

$$\begin{aligned} &\text{Minimize } \langle p, q\rangle + \gamma\sum_{i=1}^{m_1} p_i^1(\frac{1-\alpha}{\alpha}u_{1i}+v_{1i}) \\ &\text{s.t} \;\; u_{1i}-v_{1i} = q_i-\eta_1,\; i=1..m_1 \\ &q_i = -\sum_{k=1}^{m_2} p_{ik}^2 X_{2ik} + \gamma\sum_{k=1}^{m_2} p_{ik}^2(\frac{1-\alpha_2}{\alpha}u_{2ik}+v_{2ik}),\; i=1..m_1,\; k=1..m_2 \\ &\text{s.t} \;\; u_{2ik}-v_{2ik} = \eta_{2i}-X_{2i},\; i=1..m_1,\; k=1..m_2 \\ &u_{2i}\geq 0,\; v_{2i}\geq 0,\quad z_i^2\in Z_{2i},\; \eta_{2i}\in\mathbb{R},\; i=1..m_1 \\ &z^1\in Z_1,\; \eta_1\in\mathbb{R}, u_1\geq 0,\; v_1\geq 0. \end{aligned} \tag{6.43}$$

Chapter 7

Numerical Experiments (Part 2)

7.1 Objectives

The main objectives of the numerical experiments are to calculate and interpret the risk-adjusted probability measures in two stage problems, and to compare the performance of cutting plane methods. With this in mind, the numerical results are broken into three main sections.

In the first section, the two-stage mean-risk portfolio problem, for the semideviation (5.11) and weighted mean-deviation from quantile (5.33) risk functionals are solved, for different values of the risk aversion parameter γ.

The data set from which the portfolio is drawn consists of a set of 100 assets, taken from the S& P500 index. The data set includes daily returns from the last 528 days of trading for each asset. The weekly returns are constructed from daily returns, taken over 5 business days. A two-stage scenario tree of weekly returns is randomly sampled from this data set, with 50 nodes in the first stage, and 40 nodes in the second stage corresponding to each node in the first stage.

In the second section, for the risk-aversion parameter $\gamma = 0.9$ and trading cost $\kappa = 0.005$, the first- and second-stage risk-adjusted probability measures are calculated, for the semideviation (5.11) and weighted mean-deviation from quantile (5.33) risk functionals. These measures are used to measure portfolio performance in a similar fashion to the one-stage portfolio problem.

In the third section, the two-stage mean-risk portfolio problem, for the semideviation (5.11) and weighted mean-deviation from quantile (5.33) risk functionals are solved, using three different methods: formulate as a large linear program and apply simplex method; use extended Benders' decomposition method; use extended multi-cut Benders' method. The performance of the three methods are compared for different size scenario trees.

7.2 Risk Aversion Parameter and Trading Costs

7.2.1 Risk Aversion Parameter

In this section, the two-stage mean-risk portfolio problem, for semideviation (5.11) and mean-weighted deviation from quantile (5.33) risk functionals are solved, for the following values of the risk aversion parameter γ

$$\gamma = \{0.1, 0.3, 0.5, 0.9\} \tag{7.1}$$

The cumulative distribution functions (CDF) of end portfolio returns was constructed for each value of γ, using the original probability measures. These curves were plotted against each other. The objective was to examine how the risk aversion parameter affects the outcome of the portfolio.
The graphs for the comparison of different risk aversion constant are presented in Figures 7.1 and 7.2, for the semideviation and weighted deviation from quantile risk functionals, respectively. In Figure 7.1, for the semideviation risk function, it can be seen that the curves have smaller tails, as the value of γ increases. However, as the risk aversion constant γ increases, the probability of higher returns decreases.

In Figure 7.2, for the weighted deviation from quantile risk function, the curves are significantly further apart than in Figure 7.1. The curves have much shorter tails for larger values of γ.

In summary, for the two-stage mean-risk portfolio problem, using semideviation (5.11) and weighted deviation from qantile (5.33) risk functionals, the range of the curves decreases as the size of the risk aversion parameter γ increases. This pattern is much more pronounced in the mean-weighted deviation from quantile case, perhaps due to the very high penalty for returns in the left tail. This pattern reflects the fact that taking less risk decreases the chances for both very high and very low returns on investment.

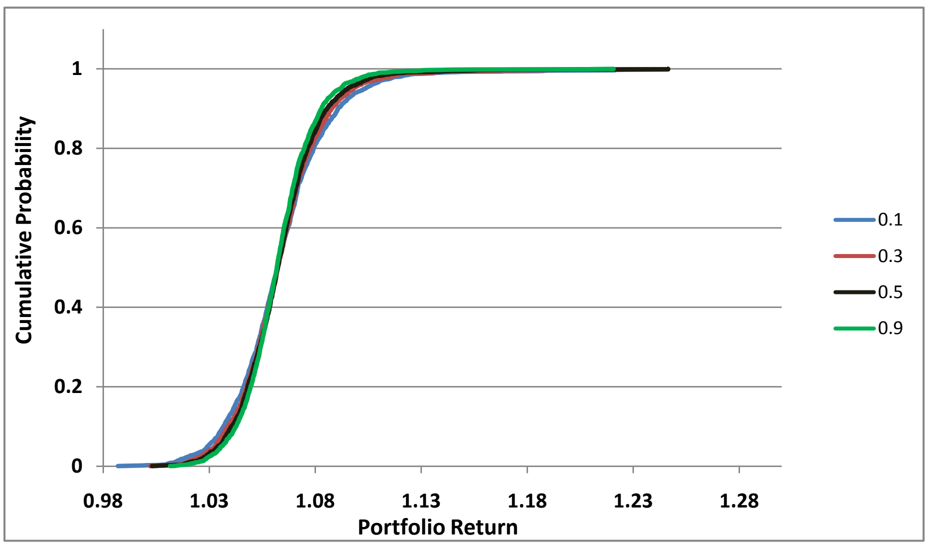

Figure 7.1: Cumulative distribution curves of the mean-semideviation optimal portfolio, for different values of the risk aversion parameter γ.

7.2.2 Trading Costs

In this section, the two-stage mean-risk portfolio problem, for semideviation (5.11) and mean-weighted deviation from quantile (5.33) risk functionals are solved, for the following values of the trading cost parameter κ

$$\kappa = \{0.1, 0.05, 0.01, 0.005\} \tag{7.2}$$

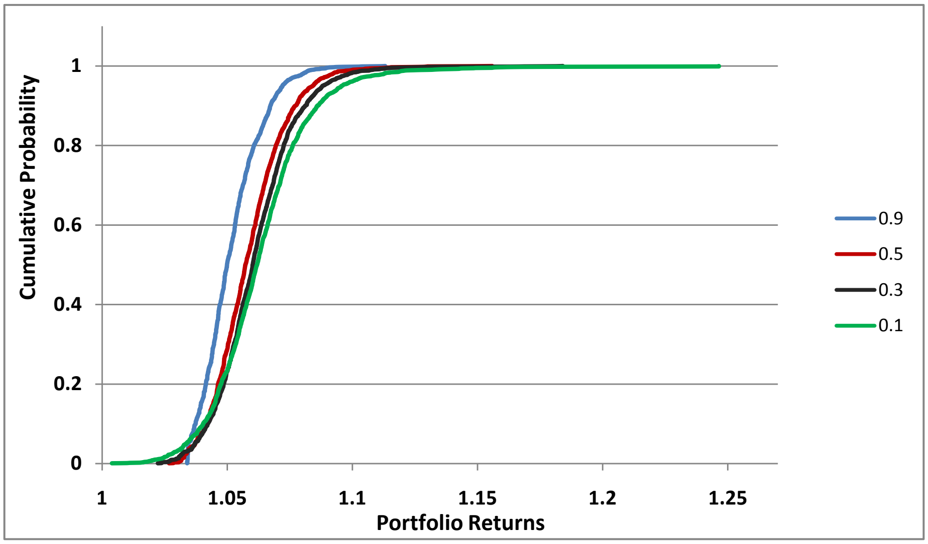

Figure 7.2: Cumulative distribution curves of the mean-deviation from quantile optimal portfolio, for different values of the risk aversion parameter γ.

For each risk functional, the CDF of portfolio returns were constructed and plotted against eachother for different values of κ. The risk-aversion parameter γ is set to 0.9.

The graphs for the comparison of different trading costs κ are presented in Figures 7.1 and 7.2, for the semideviation and weighted deviation from quantile risk functionals, respectively. In both figures, the curves with lower trading cost have longer right tails. As the trading cost increases, the likelihood of higher return decreases.

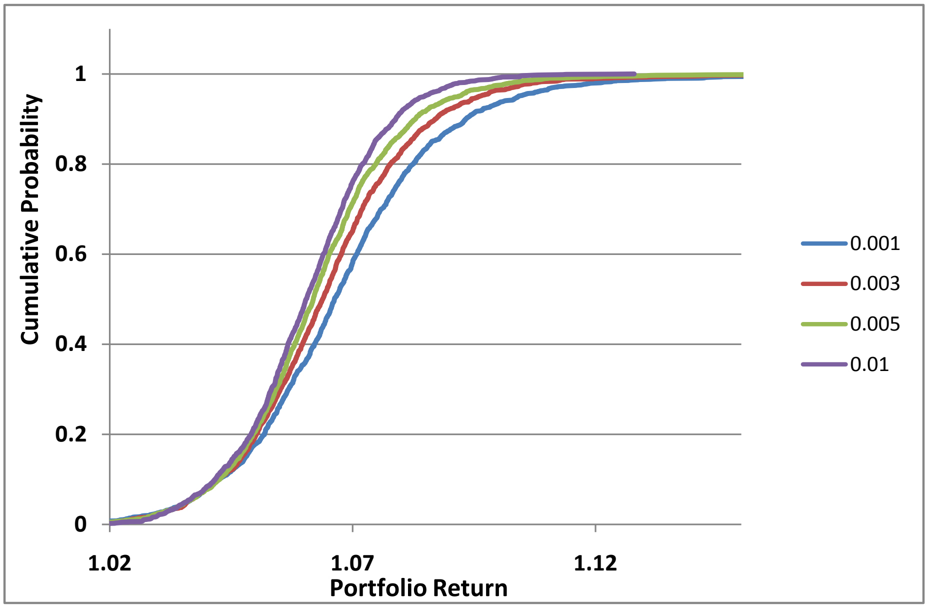

Figure 7.3: Comparison of cumulative probability distribution curves for different trading costs κ, for the semideviation risk function.

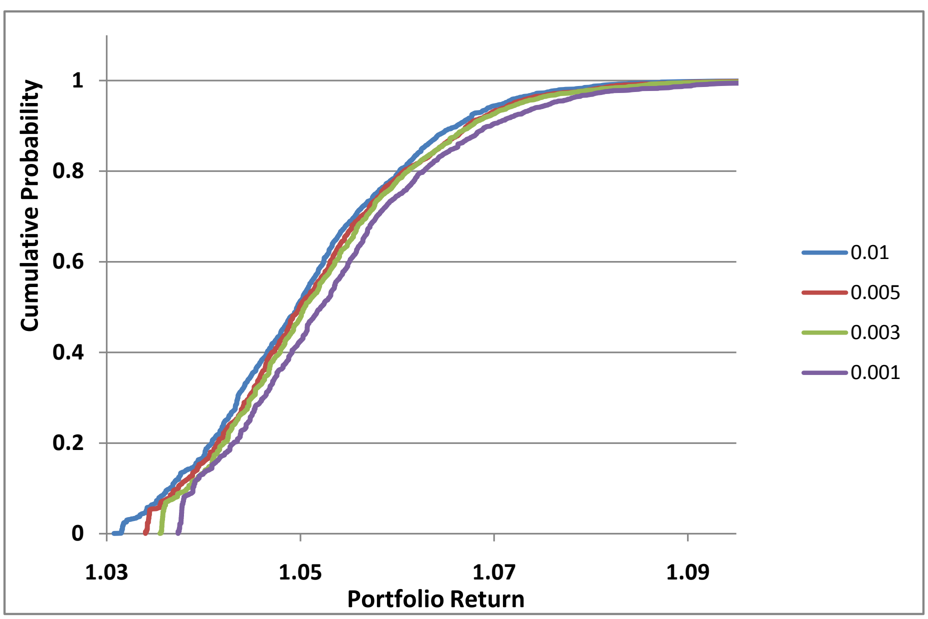

Figure 7.4: Comparison of cumulative probability distribution curves for different trading costs κ, for the mean weighted devation from quantile risk function.

7.3 Optimal Portfolios and Risk-Adjusted Probability Measures

7.3.1 Setup

In this section, the two-stage mean-risk portfolio problem (4.12) – (4.11) was solved for the risk functionals semideviation (5.11) and weighted deviation from quantile(5.33). The risk coefficient γ was set to 0.9, and trading cost κ was set to 0.005. The capital allocation was not allowed to exceed 10 percent for any asset in the portfolio.

The first and second stage risk-adjusted probability measures were calculated for the optimal portfolio, for both the mean-semideviation and mean-weighted deviation from quantile portfolio problems. The original and risk adjusted cumulative distribution functions (CDF) were constructed and plotted against each other.

Separately, a two-stage market portfolio, with each asset having equal weight, was constructed. The original and risk adjusted cumulative distribution functions (CDF) were calculated for this portfolio, and plotted against each other.

As discussed in the numerical results of the one-stage portfolio problem, plotting the CDF curves together allows us to compare the perspectives on the behaviour of the portfolio, with the risk-adjusted CDF representing the perspective of the risk-averse investor. If the curves are close together, then the optimal portfolio is robust. If the curves are far apart, then the optimal portfolio doesn't reflect the concerns of a risk-averse investor.

7.3.2 Semideviation

The optimal portfolio for the mean-semideviation portfolio problem is presented in Table 7.1. The portfolio diversity is somewhat low, despite the high risk-aversion constant. This maybe due to the restriction that at most 10 percent can be invested in any asset.

The first-stage risk-adjusted probability measures for the optimal portfolio are presented in Table 7.2. The risk-adjusted CDF and original CDF for the optimal portfolio are plotted together in Figure 7.5. The curves are far apart, suggesting the optimal portfolio does not align with the risk-averse investors preferences.

Asset	Value	Asset	Value
7	0.1	74	0.1
32	0.1	84	0.1
36	0.1	88	0.1
51	0.1	89	0.012
53	0.088	93	0.1
54	0.1	.	.

Table 7.1: The optimal two-stage mean-semideviation portfolio, Risk $= -1.05478$

1	0.0283	11	0.0283	21	0.0283	31	0.0283	41	0.0103
2	0.0283	12	0.0103	22	0.0103	32	0.0283	42	0.0283
3	0.0283	13	0.0283	23	0.0283	33	0.0283	43	0.0283
4	0.0103	14	0.0283	24	0.0103	34	0.0283	44	0.0103
5	0.0103	15	0.0103	25	0.0103	35	0.0103	45	0.0103
6	0.0103	16	0.0283	26	0.0103	36	0.0103	46	0.0283
7	0.0103	17	0.0103	27	0.0283	37	0.0283	47	0.0283
8	0.0283	18	0.0283	28	0.0283	38	0.0283	48	0.0103
9	0.0103	19	0.0103	29	0.0283	39	0.0283	49	0.0103
10	0.0103	20	0.0283	30	0.0103	40	0.0103	50	0.0283

Table 7.2: First stage risk-adjusted probability measures for the optimal two-stage mean-semideviation portfolio.

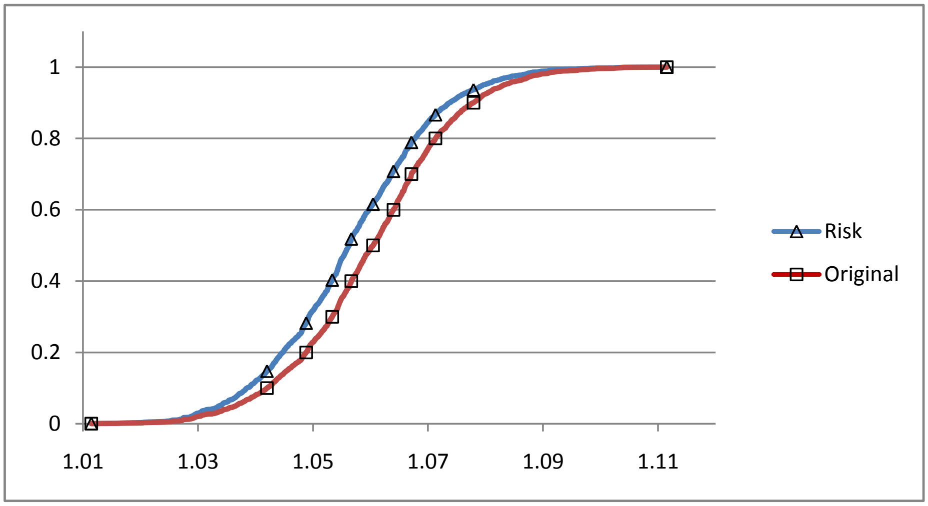

Figure 7.5: Comparison of the risk-adjusted and original cumulative probability distribution curves for the mean-semideviation optimal portfolio.

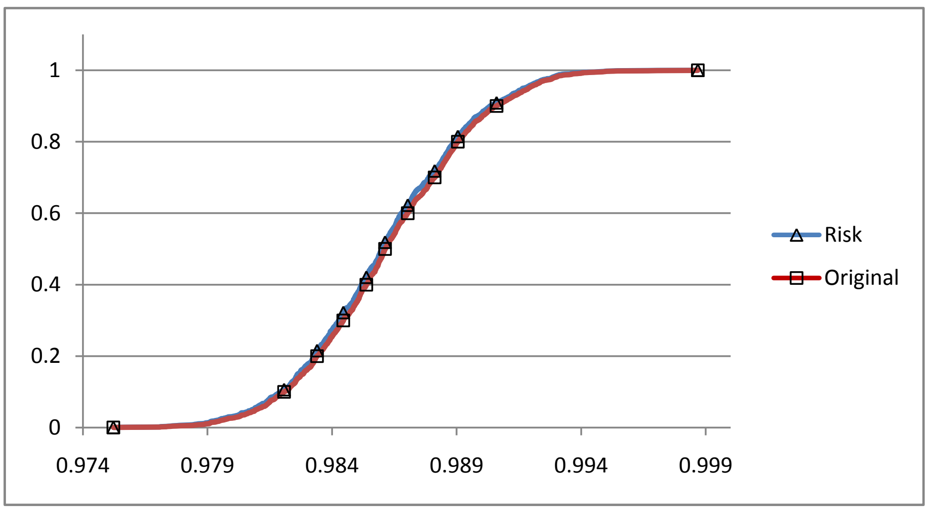

Figure 7.6: Comparison of the risk-adjusted and original cumulative probability distribution curves for the market portfolio, using the semideviation function.

The risk-adjusted CDF and original CDF for the market portfolio are plotted together in Figure (7.6), The curves are closer than in figure (7.5).

The results suggest that for the two-stage mean-semideviation portfolio problem, the optimal portfolio does not perform as well as the market portfolio, from the perspective of a risk averse investor. However, the optimal portfolio has a higher range of end portfolio values than the market portfolio.

7.3.3 Mean-Weighted Deviation from Quantile

The optimal portfolio for the mean-weighted deviation from quantile portfolio problem is presented in table (7.3). The portfolio diversity is much higher than in the semideviation case, despite the same restrictions on investments allowed in each asset. The very high penalty in the left tail may account for this.

The first-stage risk-adjusted probability measures for the optimal portfolio are presented in table (7.4). The risk-adjusted CDF and original CDF for the optimal portfolio are plotted together in Figure (7.7). The curves are far apart, suggesting the portfolio doesn't perform well from the perspective of a risk averse investor.

Asset	Value	Asset	Value
7	0.054	78	0.031
12	0.1	84	0.01
13	0.034	85	0.046
18	0.038	88	0.1
32	0.080	89	0.029
36	0.1	93	0.06
51	0.07	97	0.026
59	0.085		
74	0.048		

Table 7.3: The optimal two-stage mean-deviation from quantile portfolio, Risk = -1.03599

1	0.002	11	0.002	21	0.002	31	0.002	41	0.119
2	0.002	12	0.002	22	0.002	32	0.002	42	0.002
3	0.002	13	0.002	23	0.002	33	0.002	43	0.002
4	0.002	14	0.002	24	0.002	34	0.002	44	0.002
5	0.002	15	0.051	25	0.002	35	0.119	45	0.002
6	0.002	16	0.002	26	0.002	36	0.002	46	0.002
7	0.210	17	0.065	27	0.002	37	0.002	47	0.002
8	0.002	18	0.002	28	0.002	38	0.002	48	0.002
9	0.002	19	0.086	29	0.002	39	0.002	49	0.002
10	0.264	20	0.002	30	0.002	40	0.002	50	0.002

Table 7.4: First stage risk-adjusted probability measures for the optimal two-stage mean-deviation from quantile portfolio.

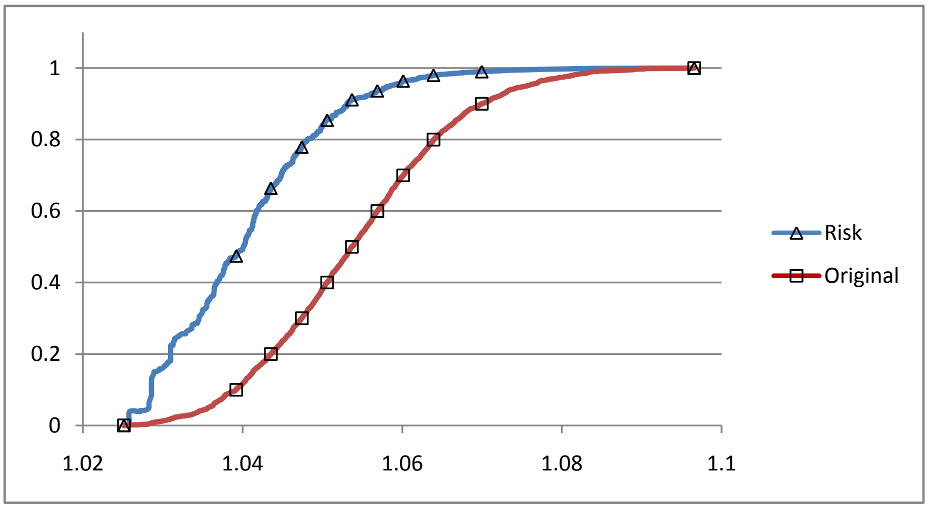

Figure 7.7: Comparison of the risk-adjusted and original cumulative probability distribution curves for the mean-deviation from quantile optimal portfolio.

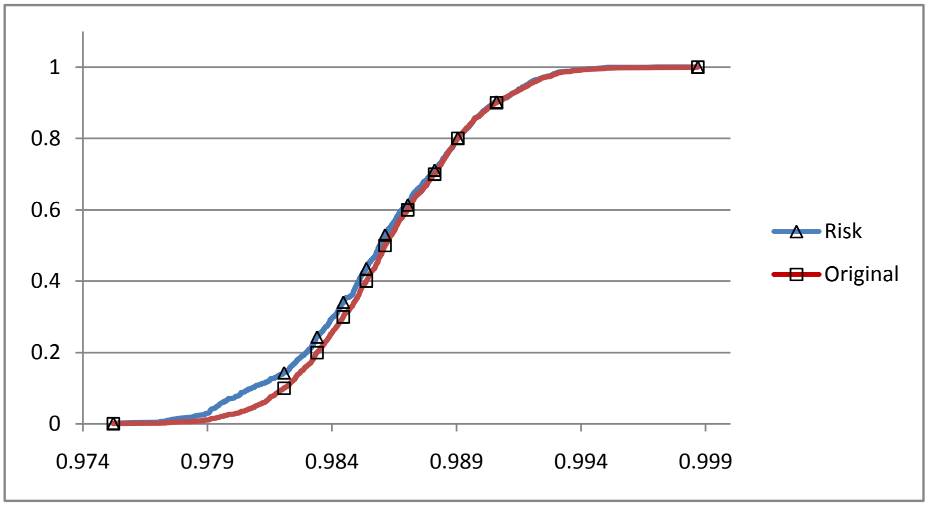

Figure 7.8: Comparison of the risk-adjusted and original cumulative probability distribution curves for the market portfolio, using the weighted mean-deviation from quantile risk function.

The risk-adjusted CDF and original CDF for the market portfolio are plotted together in Figure (7.8), were the curves are closer together than for the optimal portfolio.

The results suggest that for the two-stage mean-deviation from quantile portfolio problem, the optimal portfolio does not perform as well as the market portfolio, from the perspective of a risk averse investor. However, the optimal portfolio has a higher range of end portfolio values than the market portfolio.

7.4 Comparison of Different Solution Methods

7.4.1 Setup

In this section, we compare the performance of the three solution methods for the two-stage risk averse portfolio problem, described in (6.20), (6.31), and (6.37, 6.43). As in the previous section, a portfolio of 100 assets was drawn from the S & P500 index. Weekly returns were constructed from daily returns, taken over 5 business days. Scenario trees $\xi^1 \times \xi^2$, of different sizes were randomly sampled. Recall that ξ^1 represents the number of first stage nodes, and ξ^2 represents the number of second stage nodes connected to any first stage node. The following scenario trees were used: 10×10, 20×20, 50×40.

On each of these scenario trees, the two-stage mean-risk portfolio problem with risk measures semideviation (5.11) and weighted deviation from quantile (5.33) was constructed. The portfolio problem was solved using Benders' Decomposition method and the multi-cut Benders' Decomposition method. Separately, the portfolio problem was written as one large linear programming problem and solved using the simplex method.

The performance of the three solution methods were compared along the following dimensions: total solve time, total run time, total computer memory used, and number of simplex iterations used.

7.4.2 Results on Solve Time

The solve times, for different sized scenario trees are displayed in Tables 7.5 and 7.6, for the mean-semideviation and mean-weighted deviation from quantile portfolio problems, respectively.

Scenario	Linear	Benders	Multi
10×10	6.453	12.507	5.996
20×20	5.578	35.683	22.469
50×40	7.531	326.523	65.121

Table 7.5: Comparison of total solve time for two stage mean-semideviation portfolio problem, using different solution methods.

Scenario	Linear	Benders	Multi
10×10	6.345	10.769	8.193
20×20	8.75	28.938	22.359
50×40	6.906	235.594	166.405

Table 7.6: Comparison of total solve time for two stage mean-deviation from quantile portfolio problem, using different solution methods.

Looking at the tables, it is clear that the linear programming method outperformed the cutting plane methods, in all three scenario trees for the quantile risk measure and in all but one scenario tree for the semideviation risk measure. As the size of the scenario tree increased, this gap in run times between the linear method and the cutting plane methods increased substantially. Between the two cutting plane methods, the multi-cut Benders decomposition method generally took less time to solve than the Benders decomposition method. The gap in solve times between the two cutting plane methods was much smaller than the gap between the linear method and any of the cutting plane methods.

7.4.3 Results on Memory Usage

We begin this section with the results on the dimension of total memory used by the three methods, for the two stage mean-semideviation model. The results are displayed in Table 7.7

Scenario	Linear	Benders	Multi
10×10	2,813,800	2,559,752	2,302,736
20×20	7,697,632	6,335,928	6,908,624
50×40	31,402,560	20,914,800	27,812,564

Table 7.7: Comparison of total memory used by different solution methods, for the two stage mean-semideviation portfolio problem.

Looking at the table, it is evident that the total memory used by the linear method exceeded that used in the extended Benders' and multi-cut Benders' methods for all three scenario trees. The total memory used by the extended multi-cut Benders' method was less than that used by the extended Benders' method nad linear method for all three scenario trees.

To obtain a better perspective on these results, a table of ratios was constructed, for each scenario tree 7.8. The ratios are meant to compare the memory used by one method against another. For example, a ratio was constructed with the memory used by Benders' Decomposition in the numerator, and the memory used by the linear method in the denominator. Ratios comparing the memory usage of the extended multi-cut Benders decomposition method to the linear method, and the extended Benders decomposition method to the extended multi-cut Benders decomposition method were also constructed.

By placing these ratios together for different sized scenario trees, we can see how the memory usage depends on the size of the tree. For example, in the first case, if the ratio decreases in size as the scenario tree increases in size, then the Benders' Decomposition method memory usage grow slower in proportion to the linear programs' memory usage. For very large programs, this may mean the Linear method will be more likely to crash where the extended Benders' decomposition method will solve. The table of proportions is given below:

Scenario	$\frac{\text{Benders memory}}{\text{Linear memory}}$	$\frac{\text{Multi Cut Benders memory}}{\text{Linear memory}}$	$\frac{\text{Benders memory}}{\text{Multi Cut Benders memory}}$
10×10	0.9097	0.8184	1.1116
20×20	0.8231	0.8975	0.9171
50×40	0.666	0.8857	0.7520

Table 7.8: Comparison of ratios for memory usage, for different solution methods, for two stage mean semideviation portfolio problem.

Looking at Table (7.8), it is clear that the proportion of memory used by both the extended Benders' decomposition method and the extended multi-cut Benders' decomposition method, in relation to the linear method are less than one, suggesting they use less memory than the linear method in all three scenario trees. The proportions decrease as the scenario tree increases in size for the Benders method comparison, and stays between 0.8 and 0.9 for the multi cut Benders comparisons. For a scenario tree of size 500×200, we tested both the linear and extended Benders decomposition methods. The linear programming method crashed because of memory overload, while, albeit very slowly, the extended Benders' decomposition method solved.

For the mean-deviation from quantile model, the total memory used by the three methods is presented in Table 7.9.

The linear model used more memory than the extended multi-cut Benders' decomposition method on all three scenario trees, and used less memory than the

Scenario	Linear	Benders	Multi
10 × 10	2,906,240	2,664,048	2,296,880
20 × 20	7,869,800	8,669,176	6,085,064
50 × 40	32,220,472	39,482,832	31,543,752

Table 7.9: Comparison of total memory used by different solution methods, for the two stage mean-deviation from quantile portfolio problem.

extended Benders' decomposition method on all but one scenario tree.

As in the semideviation case, a table of proportions was constructed. The table of these proportions is given below :

Scenario	$\frac{\text{Benders memory}}{\text{Linear memory}}$	$\frac{\text{Multi Cut Benders memory}}{\text{Linear memory}}$	$\frac{\text{Benders memory}}{\text{Multi Cut Benders memory}}$
10 × 10	0.917	0.7903	1.1599
20 × 20	1.10	0.7732	1.4247
50 × 40	1.225	0.9790	1.2517

Table 7.10: Comparison of ratios for memory usage, for different solution methods, for the two stage mean deviation from quantile portfolio problem.

In the first column of Table (7.10), the ratio of memory used by extended Benders method in relation to the linear method increases in the size of the scenario tree,albeit somewhat slowly. The ratios in the second column, representing comparison of memory usage by extended multi-cut Benders' decomposition method to the linear method, increase somewhat slowly as the scenario tree size increases. The ratio of memory used by extended Benders' method in relation to the extended multi-cut Benders' method, in the third column, stays in the range of 1.15 to 1.43.

We note that in both cases, the computer programs to solve the two stage mean risk portfolio problems did not use the most efficient storage techniques for memory. The optimal portfolio, first stage risk-adjusted probability measure and Lagrange multipliers were recorded and stored for each iteration. With improved memory storage techniques, these ratios may change.

7.4.4 Results on Number of Iterations

The number of outer iterations for the extended Benders and the multi-cut Benders decomposition methods are presented in Tables 7.11 and 7.12, for the mean-semideviation and mean-weighted deviation from quantile models, respectively.

Scenario	Benders	Multi
10 × 10	18	7
20 × 20	22	10
50 × 40	19	14

Table 7.11: The number of outer iterations generated in the master problem for the two cutting plane methods, for the two-stage mean-semideviation portfolio problem.

Scenario	Benders	Multi
10 × 10	20	10
20 × 20	37	14
50 × 40	81	35

Table 7.12: The number of outer iterations generated in the master problem for the two cutting plane methods, for the two-stage mean-deviation from quantile portfolio problem.

The number of outer iterations in the mean-semideviation model did not increase substantially over the different scenario trees. The extended Benders' method used a larger number of outer iterations than the multi-cut Benders' decomposition methods at each scenario tree. For the mean-weighted deviation from quantile model, the pattern was repeated. However, the number of outer iterations increased susbstantially, as the size of the scenario tree increased.

7.5 Comparison of Aggregate and Conditional Risk Mapping Approach

In this section, we construct the two-stage mean-semideviation and mean-deviation from quantile portfolio problems, with 50 outcomes in the first stage, and 40 second stage outcomes for each first stage outcome. For the trading costs $\kappa = 0.005$ and $\kappa = 1$, we compare the end portfolio return outcomes. The second problem is equivalent to a one stage model with no trading. The question we are addressing is whether or not it makes sense to trade.

In the two-stage mean-risk porfolio problem with risk function semideviation, there was not much difference in the portfolio diversification between the different trading costs. This may be due to the restriction of 10 percent or less being invested in any asset. The optimal objective was better with the lower trading cost.

Asset	Value
7	0.1
32	0.1
36	0.1
51	0.1
53	0.088
54	0.1
74	0.1
84	0.1
88	0.1
89	0.012
93	0.1

Table 7.13: Optimal portfolio for the two-stage mean-semideviation portfolio problem, with trading cost 0.005. Optimal objective value Risk $= -1.05478$

Asset	Value
7	0.1
32	0.1
36	0.1
51	0.1
53	0.071
54	0.1
74	0.1
84	0.1
85	0.028
88	0.1
93	0.1
97	0.0002

Table 7.14: Optimal portfolio for the two-stage mean-semideviation portfolio problem, with trading cost 1. Optimal objective value Risk $= -1.05331$

In the two-stage mean-risk porfolio problem with risk measure mean weighted deviation from quantile, again there was not much difference in the portfolio diversification between the different trading costs. This may be due to the restriction of 10 percent or less being invested in any asset. The optimal objective was better with the lower trading cost.

Asset	Value		
7	0.054	78	0.031
12	0.1	84	0.01
13	0.034	85	0.046
18	0.038	88	0.1
32	0.080	89	0.029
36	0.1	93	0.06
51	0.07	97	0.026
59	0.085	100	0.00
74	0.048		

Table 7.15: Optimal portfolio for the two-stage mean-deviation from quantile portfolio problem, with trading cost 0.005. Optimal objective value, Risk $= -1.03599$

Asset	Value		
7	0.052	12	0.050
13	0.027	18	0.028
32	0.086	36	0.1
51	0.076	53	0.003
54	0.007	59	0.059
74	0.1	84	0.0939
85	0.068	88	0.1
89	0.051	93	0.088
97	0.011		

Table 7.16: Optimal portfolio for the two-stage mean-deviation from quantile portfolio problem, with trading cost 1. Optimal objective value Risk $= -1.03079$

7.6 Progress of Benders' Decomposition

In this section, we analyze the performance of the extended Benders' Decomposition method on the two stage mean-risk portfolio problem, with risk function semideviation. For a very large scenario tree (100×100), a graph of the gap between the optimal solution at the k^{th} outer iteration, and the value of the k^{th} optimality cut function was constructed. Recall that in the Benders' Decomposition method, the algorithm stops when this gap is smaller than some ϵ, were $\epsilon > 0$. The objective was to determine how man iterations it would take to reach an optimal solution, and how fast the method converges.

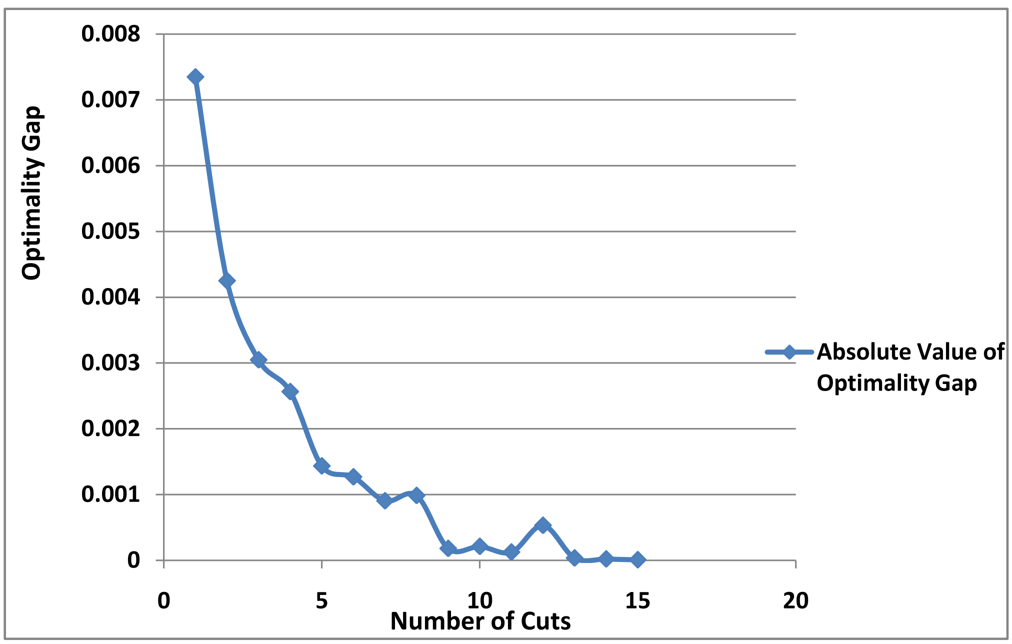

Figure 7.9: Graph of optimality gap to outer iteration number, for Benders decomposition method, applied to mean-semideviation two-stage portfolio problem.

In figure 7.9, the method converges rapidly within the first 10 to 13 outer iterations. After this, it will run for a very long time, making very small contributions to the optimal solution. For larger scenario trees, this tail grows longer, and the method will take a long time to solve.

Chapter 8

Conclusion

We formulated the one stage risk averse portfolio problem (2.3) as a zero sum matrix game, with a new set of risk-adjusted probability measures part of the optimal saddle point solution. Closed form solutions for the risk adjusted probability measures were constructed, for the risk averse portfolio problem with mean-risk objective, for the risk functions semideviation (5.10) and mean-weighted deviation from quantile (5.31). These measures were used to evaluate the perfomance of the optimal portfolio, compared to a market portfolio. The results suggest that the optimal portfolio performs better than the market portfolio.

In the second part of the dissertation, we introduced the conditional risk mapping approach to the problem of optimizing a portfolio over two investment periods, with the option to rebalance. The resulting two-stage portfolio optimization problem was called the two-stage risk-averse portfolio problem (4.12).

Using convex analysis, the two-stage risk averse portfolio problem was reformulated as a zero sum matrix game, with first stage risk adjusted probability measures derived from the optimal saddle point solution. The second stage problem was also formulated as a zero-sum matrix game, and optimal risk adjusted probability measures were calculated.

The first and second stage risk adjusted probability measures were calculated for the mean-risk portfolio problem, for the risk functions semideviation (5.10) and mean-weighted deviation from quantile (5.31). As in the first section, these measures were used to evaluate the performance of the optimal portfolio, compared to a market portfolio. The results suggest the optimal portfolio performs better.

The Benders' Decomposition and Multi-cut Benders' Decomposition methods for solving a two-stage linear stochastic programming problem (6.2) were extended to solving the two-stage risk-averse portfolio problem. On scenario trees of different sizes, the performance of these methods and a linear programming method were compared on the specs of total solve time, and memory usage, for both the mean-semideviation and mean-weighted deviation from quantile models. On the specs of time, the linear program performed better on all scenario trees. For total memory usage, the hypothesis was that the extended Benders' methods would use less memory in relation to the linear program as the scenario tree size increased. This hypothesis was supported for the mean-semideviation model. On very large scenario trees, it crashed were the extended Benders' method solved. Another hypothesis, that the extended multi-cut Benders' decomposition method would use less memory than the extended Benders' decomposition method, was not supported for either the mean-semideviation or the mean-weighted deviation from quantile models.

It was mentioned that the computer programs for the extended Benders' decomposition methods did not make use of the most efficient memory storage methods. In the future, these programs could be run with better memory storage techniques, to yield improvements on the performance of the extended Benders' decomposition methods.

Also, for future research, it would be interesting to examine the question of which risk measures to use in each stage of the two stage risk averse portfolio problem. Recall we used the mean-semideviation risk measures for both objective functions in one model, and mean deviation from quantile risk measures for both objectives in another model. This was done for illustrative purposes, to examine how risk adjusted probability measures could evaluate portfolio performance, and to test the extensions on Benders' decomposition methods. The question of what the composition of these two measures means intuitively, or wether other compositions would be more useful would be interesting to investigate further.

References

[1] Artzner, P., F.Delbaen, J.M.Eber and D.Heath, Thinking Coherently, Risk, 10, 68-71 1997.

[2] Artzner, P., F.Delbaen, J.M.Eber and D.Heath, Coherent measures of risk, Mathematical Finance, 9, 203-228, 1999.

[3] Artzner, P., F.Delbaen, J.M.Eber, D.Heath, and H. Ku, Multiperiod Risk and Coherent Multiperiod Risk Measurement, 2002.

[4] Acerbi C., Tasche D., Expected Shortfall: A natural coherent alternative to Value at Risk, Economic Notes, volume 31, issue 2, 379-388, 2003.

[5] Basel Committee on Banking Supervision, The new Basel Capital Accord, Basel Committee on Banking Supervision, Consulative Document, BIS January 2001.

[6] Brachinger, H. W., and Weber, M., Risk as a primitive: a survey of measures of percieved risk, Operations Research- Spektrum, 19(1997), 235-250.

[7] Domar E.V., and Musgrave R.A., Proportional Income Taxation and Risk Taking, Quart. J. Econ., 58, 389-422, May 1944.

[8] Duffie D., and J. Pan, An Overview of the Value at Risk, Journal of Derivatives, 4, 7 - 49, 1997.

[9] Fishburn, PC, Mean-Risk Analysis with Risk Associated with Below-Target Returns, The American Economic Review, 67, 116-126, 1977.

[10] Föllmer, H., and A. Schied, Convex measures of risk and trading constraints, Finance and Stochastics, 6, 429-447, 2002.

[11] Föllmer, H., and A. Schied, Stochastic Finance. An Introduction in Discrete Time, de Gruyter, Berlin, 2004.

[12] Fritelli, M, and E. Rosazza Gianin, Putting order in risk measures, Journal of Banking and Finance, 26, 1473-1486, 2002.

[13] Hanoch, G., and H., Levy, The Efficiency Analysis of Chocies Involving Risk , Review of Economic Studies, 36, 335-346, 1969.

[14] Hardy G.H., J.E. Littlewood, and G. Polya, Inequalities, Cambridge U.K.: Cambridge University Press, 1934.

[15] Klein Haneveld, W.K., Duality in Stochastic Linear and Dynamic Programming, Lecture Notes in Economics and Math. Systems 274, Springer, Heidelberg, 1986.

[16] Kusuoka, S., On Law Invariant Coherent Risk Measures, Adv. Math. Econ., 3, 83-95, 2001.

[17] Mao J.C.T., Survey of Capital Budgeting: Theory and Practice, J. Finance, 25, 349-360, 1970.

[18] Markowitz, H.M. (1952): Portfolio Selection, Journal of Finance, 7, 77-91

[19] Markowitz, H.M. Portfolio Selection: Efficient Diversification of Investment, Yale University Press, New Haven, USA, 1959.

[20] Mansini, R., Ogryczak, W., and M. G. Speranza, On LP solvable models for portfolio selection, Informatica, 14 (1), 37-62, 2003.

[21] Miller N. , Ruszczyński, A., Risk-Adjusted Probabilities in Portfolio Optimization with Coherent Risk Measures, European Journal of Operational Research, vol 191, issue 1, 193-206, 2008.

[22] Ogryczak, W., Ruszczyński, A., From Stochastic Dominance to mean-risk models: Semideviations as risk measures, European Journal of Operational Research, 116(1999) 33-50

[23] Ogryczak, W., Ruszczyński,A., On Consistency of Stochastic Dominance and Mean-Semideviation Models, Mathematical Programming, 89, 217-232, 2001.

[24] Ogryczak, W., Ruszczyński, A., Dual Stochastic Dominance and Related Mean-Risk Models, SIAM J. on Optimization, 13 (2002), 60-78.

[25] Porter, R.B., Semivariance and Stochastic Dominance: A comparison, The American Economic Review 64, 200-204, 1974.

[26] Porter, R.B., Gaumnitz, J.E., Stochastic Dominance v.s. Mean-Variance Portfolio Analysis: An Empirical Evaluation, The American Economic Review, 62, 438-446, 1972.

[27] Prékopa A., Stochastic Programming, Kluwer Academic Publishers, London 1995.

[28] Prékopa A., Contributions to the Theory of Stochastic Programming, Math. Prog. 4, 202-221, 1973.

[29] Riedel, F., Dynamic Coherent Risk Measures, Stochastic Processes Appl., 112, 185-200, 2004.

[30] Rockafellar R. T., Conjugate Duality and Optimization, SIAM, 1987.

[31] Rockafellar R.T, Uryasev, S., and M. Zabarankin, Deviation measures in risk analysis and optimization, Finance and Stochastics, 10, 51-74, 2006.

[32] Rockafellar R.T, Uryasev, S., Optimization of Conditional Value at Risk, The Journal of Risk, 2, 21-41, 2000.

[33] Rothschild, M., Stiglitz, J.E., Increasing Risk: 1 A Definition, Journal of Economic Theory 2, 225-243, 1970.

[34] Ruszczyński, A, Nonlinear Optimization, Princeton University Press, Princeton 2006.

[35] Ruszczyński, A. and A. Shapiro, Optimization of convex risk functions, Mathematics of Operations Research, 31, 433-452, 2006.

[36] Ruszczyński A. and Shapiro A., Optimization of risk measures, In Probabilistic and Randomized Methods for Design under Uncertainty, G. Calafiore and F. Dabbene, (Eds.), Springer, London, 2005

[37] Ruszczyński, A and Shapiro A., Conditional Risk Mappings, Mathematics of Operations Research, Vol 31, No. 3, August 2006, pp 544-561

[38] Ruszczyński, A. and R. J. Vanderbei, Frontiers of Stochastically Nondominated Portfolios, Econometrica, 71(4), 1287-1297, 2003.

[39] Ruszczyński, A. and Shapiro A., Optimality and Duality in Stochastic Programming,

[40] Shapiro A. and Ruszczyński, A, Lectures on Stochastic Programming, November 19 2007.

[41] Cheng S., Liu Y., and Wang S., Progress in Risk Measurement, Advanced Modelling and Optimization, Volume 6, Number 1, 2004.

[42] Stone B.K., A General Class of Three Parameter Risk Measures, Journal of Finance, 28, 675-685, 1973.

[43] Tasche D., Expected Shortfall and Beyond, Journal of Banking and Finance, 26, 1519-1533, 2002.

[44] Tobin J., Liquidity Preferences as a Behaviour Toward Risk, Review of Economic Studies, 25, 65-86, 1958.

[45] Whitmore, G.A., Mc. (Eds.), Stochastic Dominance: An Approach to Decision-Making Under Risk, Heath, Lexington M.A. 1978.

[46] Yamai Y. and Yoshiba T., Comparitive analyses of expected shortfall and Value-at-Risk under Market Stress, Monetary and Economic Studies, volume 20, 3, 181-237, 2002.

[47] Yitzhaki S., Stochastic Dominance, Mean Variance and Gini's Mean-Difference, The American Economic Review, 72, 178-185, 1982.

Vita

Naomi Miller

2004-2008 Ph.D. in Operations Research, Rutgers University

1999-2003 B. Sc. (Hons.) Mathematics, University of Toronto

2007-2008 Teaching Assistant Rutgers University, 1st year Calculus and Pre-calculus

2006-2007 Research Assistant, Rutgers University

2008 Miller N. , Ruszczyński, A., Risk-Adjusted Probabilities in Portfolio Optimization with Coherent Risk Measures, European Journal of Operational Research, vol 191, issue 1, 193-206,