MULTICRITERIA PORTFOLIO OPTIMIZATION

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

CEREN TUNCER ŞAKAR

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
INDUSTRIAL ENGINEERING

SEPTEMBER 2013

Approval of the thesis

MULTICRITERIA PORTFOLIO OPTIMIZATION

Submitted by **CEREN TUNCER ŞAKAR** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Industrial Engineering Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen ____________________
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Murat Köksalan ____________________
Head of Department, **Industrial Engineering**

Prof. Dr. Murat Köksalan ____________________
Supervisor, **Industrial Engineering Dept., METU**

Examining Committee Members

Assoc. Prof. Dr. Yasemin Serin ____________________
Industrial Engineering Dept., METU

Prof. Dr. Murat Köksalan ____________________
Industrial Engineering Dept., METU

Prof. Dr. Nuray Güner ____________________
Business Administration Dept., METU

Assoc. Prof. Dr. Esra Karasakal ____________________
Industrial Engineering Dept., METU

Assist. Prof. Dr. Banu Yüksel Özkaya ____________________
Industrial Engineering Dept., Hacettepe University

Date: 26.09.2013

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: Ceren Tuncer Şakar

Signature:

ABSTRACT

MULTICRITERIA PORTFOLIO OPTIMIZATION

Tuncer Şakar, Ceren
Ph.D., Department of Industrial Engineering
Supervisor: Prof. Dr. Murat Köksalan

September 2013, 144 pages

Portfolio optimization is the problem of allocating funds between available investment options in the financial market. This thesis develops several approaches to multicriteria portfolio optimization. The use of multiple criteria is justified by demonstrating their effects on decision and objective spaces of the problem. The performance of a genetic algorithm with two and three criteria is studied; and a preference-based genetic algorithm to solve portfolio optimization with complicating constraints is developed. Furthermore, stochastic programming is used to handle multi-period problems, and several issues are studied with this approach. Efficient market hypotheses, random walk and single index models are discussed in the context of scenario generation for the Turkish Stock Market. An interactive approach to stochastic programming-based portfolio optimization is also developed to guide the decision maker toward preferred solutions. The approaches are experimented with and demonstrated using stocks from the Turkish Stock Market.

Keywords: Portfolio optimization, multiple criteria decision making, genetic algorithms, stochastic programming.

ÖZ

ÇOK KRİTERLİ PORTFOLYO OPTİMİZASYONU

Tuncer Şakar, Ceren
Doktora, Endüstri Mühendisliği Bölümü
Tez Yöneticisi: Prof. Dr. Murat Köksalan

Eylül 2013, 144 sayfa

Portfolyo optimizasyonu mevcut fonların finans piyasasındaki yatırım seçenekleri arasında dağıtılması problemidir. Bu tez çok kriterli portfolyo optimizasyonuna çeşitli yaklaşımlar getirmektedir. Çok kriter kullanılmasının yararları kriterlerin problemin karar ve amaç uzaylarındaki etkilerinin araştırılmasıyla gösterilmiştir. Bir genetik algoritmanın performansı iki ve üç kriterle incelenmiş ve kısıtlarla zorlaştırılmış portfolyo optimizasyonu için karar verici tercihi tabanlı bir genetik algoritma geliştirilmiştir. Ek olarak, stokastik programlama çok dönemli portfolyo optimizasyonu için kullanılmış ve bu yaklaşım üzerine çeşitli çalışmalar yapılmıştır. Türk Hisse Senedi Piyasası için senaryo üretilmesi kapsamında etkin piyasa hipotezleri, rastgele seyir ve tek endeks modelleri tartışılmıştır. Ayrıca stokastik programlama tabanlı portfolyo optimizasyonunda karar vericiyi tercih ettiği çözümlere yönlendirmek için etkileşimli bir yaklaşım geliştirilmiştir. Önerilen yaklaşımlar Türk Hisse Senedi Piyasası'ndan hisse senetleriyle denenmiştir.

Anahtar Kelimeler: Portfolyo optimizasyonu, çok kriterli karar alma, genetik algoritmalar, stokastik programlama.

to Deniz

ACKNOWLEDGMENTS

My deepest gratitude goes to my supervisor Prof. Dr. Murat Köksalan for his insight, encouragements and advice throughout the research. I would like to express my sincerest appreciation for his support and guidance during my study in Industrial Engineering.

I would like to thank examining committee members Assoc. Prof. Dr. Yasemin Serin, Prof. Dr. Nuray Güner, Assoc. Prof. Dr. Esra Karasakal, Assist. Prof. Dr. Banu Yüksel Özkaya, and also Dr. Emre Balıbek for their valuable criticism and contributions.

I also wish to thank TÜBİTAK for financially supporting my academic work.

I am forever thankful to my family and friends for their love and support. Last, but definitely not least, I am grateful to my baby girl for the love and joy she brought to my life.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

BIST: Borsa Istanbul

CVaR: Conditional Value at Risk

DM: Decision Maker

ISE: Istanbul Stock Exchange

MAD: Mean Absolute Deviation

MCDM: Multiple Criteria Decision Making

MUA: Mean Under-Achievement

NSGA-II: Nondominated Sorting Genetic Algorithm II

PO: Portfolio Optimization

SP: Stochastic Programming

TSM: Turkish Stock Market

VaR: Value at Risk

CHAPTER 1

INTRODUCTION

Portfolio optimization (PO) is the problem of choosing between available investment instruments in the financial market. Examples of these instruments are stocks, bonds, mutual funds, options and deposit accounts. The decision maker (DM) of the problem, the investor, may be an individual or an institution. The primary goal is to maximize the wealth resulting from the investment; however, the DM can face several complicating conditions and constraints. The primary aspect that complicates the problem is the uncertainty involved in the progression of economical factors. Because of this uncertainty, PO at least needs to account for the risk involved in the outcome of the investment. As a result, basic models of PO have two objectives: maximizing return and minimizing risk. Several measures of return and risk have been proposed and experimented with. In particular, numerous approaches to define and model risk have been developed.

As the literature on PO expanded, needs of unconventional investors had to be taken into account. Investors may have additional concerns beside return and risk. They may require that their portfolio considers certain aspects of different kinds; such as liquidity, turnover, intermediary payments and social responsibility. In addition, risk of the portfolio can be modeled with multiple measures. Multiple Criteria Decision Making (MCDM) has therefore been a valuable tool in developing PO theory. Besides multiple criteria, PO theory has also welcomed several constraints to ensure that the portfolio meets certain requirements. Examples are requirements on the size of the portfolio, transaction amounts and costs. The inclusion of multiple criteria and constraints made the PO problem complex; and led some researchers to develop heuristic approaches.

Although the preliminary PO theory considered a single-period optimization setting, approaches that provide the DM with investment decisions for multiple future periods have also been proposed. These methods make use of mechanisms to model uncertain movements of financial markets, and produce decisions that cover the planning horizon of the DM.

This thesis addresses several approaches to multicriteria PO. One objective is to justify the use of multiple criteria in PO, and demonstrate their effect on decision and objective spaces of the problem. We apply different criteria, including several measures of risk to PO. We also experiment with constraints on the number of assets used in the portfolio and weights they can assume. We study the performance of genetic algorithm heuristics in PO. Another objective is to propose a method to handle multi-period problems. We study several issues related to the proposed method in detail. Including DM preferences in the investment process is another objective. Mechanisms to present the DM a single efficient solution according to her/his preferences are sought. In brief, in this thesis we cover several approaches to multicriteria PO; and consider different criteria, constraints, multi-period settings and DM preferences while doing so. We demonstrate the approaches with stocks from the Turkish Stock Market. The stock exchange entity of Turkey at the time of the experimental studies of this thesis was Istanbul Stock Exchange (ISE). Towards the end of our studies, on April 3, 2013, Borsa Istanbul (BIST) was established as the sole exchange entity of Turkey that also includes the former ISE. In this thesis, we use ISE to refer to the Turkish stock exchange entity; but the reader should note that the current abbreviation is BIST.

The thesis is organized as follows: In Chapter 2, we introduce the PO problem. General characteristics of the problem are discussed and a general literature review is provided. Since this thesis is a work of several approaches to PO, we review the literature specific to the approaches in the related chapters. We also define the criteria used throughout the thesis in Chapter 2. Different combinations of these will be used in different approaches. Chapter 3 preludes our main work by justifying the use of and demonstrating the effects of multiple criteria in PO. The effects of constraints on portfolio compositions are also studied. We cover genetic algorithm approaches in Chapter 4. First, we apply a well-known genetic algorithm to PO and discuss the results. Then, we introduce our genetic algorithm to handle DM preferences-driven PO with constraints and provide test results. Chapter 5 covers our approach to multi-period PO. We make use of Stochastic Programming (SP) to model our problem. Efficient market hypotheses, random walk models and a single index model are utilized with SP to generate scenarios to represent the progress of the financial market. We discuss results of experiments in detail. We also consider the utilization of rolling horizon settings in SP. After certain financial factors are realized, we consider updating scenario trees and revising decisions. In addition, we generate theorems, corollaries and remarks on the properties of optimal SP solutions with different criteria and time periods. In Chapter 6, we introduce an interactive approach to SP-based PO. Making use of weighted Tchebycheff programs, the proposed algorithm converges to preferred solutions according to DM preferences. It also uses confidence regions of objective function values to account for the randomness in scenarios. In Chapter 7, we conclude the thesis and discuss future work.

CHAPTER 2

THE PORTFOLIO OPTIMIZATION PROBLEM

PO is concerned with choosing assets from the financial market to invest available funds in. The return on the investment is the primary concern; however, since markets exhibit uncertain behavior, one needs to account for risk too. Following the pioneering work of Markowitz (1959), Modern Portfolio Theory has emerged and it has been studied extensively. The classical mean-variance portfolio model has two objectives: maximizing expected portfolio return in terms of weighted expected returns of assets; and minimizing portfolio risk in terms of variance. The weight of each asset is the proportion of that asset in the overall portfolio. The risk objective is composed of covariances of returns of pairs of assets; and therefore is quadratic.

The classical mean-variance portfolio model is given by:

$$\text{Max } z_1 = \sum_{i=1}^{n} E(r_i)\, x_i \tag{2.1}$$

$$\text{Min } z_2 = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i\, \sigma_{ij}\, x_j \tag{2.2}$$

$$\text{s.t.} \quad \sum_{i=1}^{n} x_i = 1 \tag{2.3}$$

where r_i is the return and $E(r_i)$ is the expected return of asset i, x_i is the proportion of asset i in the portfolio and σ_{ij} is the covariance of the returns of assets i and j.

The proportions of assets may be negative, which corresponds to shortselling. Shortselling is the act of selling the assets of another investor now and collecting the money yourself, while ensuring the original owner that you will return the same assets in the future. Shortselling is an option when the investor believes that the price of some asset will drop in the future. We do not consider shortselling in this thesis; proportions of assets are required to be nonnegative throughout the thesis.

As assets, we utilize stocks in all of our applications except for our exploratory tests to see the effects of fixed-income assets on the solutions. In our applications that use real-life data, we use stocks traded on ISE.

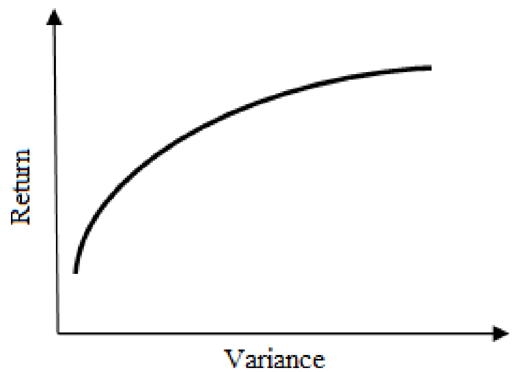


Figure 1. A typical efficient frontier for the mean-variance PO problem

Figure 1 illustrates the typical shape of the efficient frontier of the mean-variance model; it is the shape of a half bullet. Expected return increases at a decreasing rate as risk (in terms of variance) increases. The leftmost point corresponds to the minimum variance portfolio and it typically consists of several assets. Unless there is perfect positive correlation between all assets, this will be the case. In the special case of perfect positive correlation between all assets, the minimum variance portfolio will contain a single asset, the one that has the minimum variance. If there are multiple assets with the same minimum variance among these perfectly positively correlated assets, the one with the maximum expected return will be chosen. The rightmost point in Figure 1 corresponds to the maximum return portfolio and it typically consists of a single asset, the asset with the maximum expected return. In the case of multiple assets with the same maximum return value, unless they are all perfectly positively correlated, we will have a combination of these as the maximum return point in order to improve the risk objective while keeping the return objective at its best value. If all assets with the maximum return value are perfectly positively correlated, then we will select the one with minimum variance.

The shape of the efficient frontier can show some variations depending on the characteristics of the expected returns and covariances of assets. For example, if we observe a pattern where assets with high expected returns have low variances, we will obtain a squeezed efficient frontier. The two criteria will not be highly conflicting. On the contrary, if high expected returns are accompanied by high variances, the conflict will be strengthened, and the efficient frontier will be enlarged. As another case, expected returns of assets may be very close to each other. This will result in an efficient frontier that resembles a horizontal straight line with little disturbance around a certain return value. We can also experiment with different covariance matrix structures. Covariance terms are observed to be significantly smaller than variance terms in general. If we change this structure and make tests with high covariance values, we can see that the efficient frontier starts to resemble a vertical line positioned around similar variance values. Such tests can be done by utilizing random valid covariance matrices generated by the method proposed by Hirschberger et al. (2007). This method allows achieving desired distributional characteristics in the randomly generated matrix.

Working with discrete solutions on the efficient frontier, as we move from one portfolio to a neighbor, we generally observe that the same assets change their proportions up to a certain point. After that point, typically some assets are replaced and a similar trend is repeated.

2.2 Portfolio Optimization in the Literature

Looking at the early literature on PO, we see a number of different approaches after Markowitz's initial parametric optimization approach to (2.1)–(2.3). The earlier concern among researchers was to make the risk objective easier to handle. An initial approach was to diagonalize the covariance matrix to make it less dense, and thus make the problem easier to solve. Another approach was to use linear risk measures. Konno and Yamazaki (1991) proposed the use of Mean Absolute Deviation (MAD) instead of variance as a linear measure of risk. In this measure, each asset's risk is evaluated based on its return's deviation from its mean for each time point. The idea of considering only downside deviations and penalizing bigger ones harsher also emerged. Konno (1990) proposed an extension to MAD that assigns additional deviations to points that fall below some target rate of return. Michalowski and Ogryczak (2001) proposed another similar extension with multiple target levels with increasing penalties. With time, various other linear risk measures for PO were also studied. Mansini et al. (2003) provided an overview of linear programming-solvable models for portfolio selection. They examined several measures such as MAD extensions, Conditional Value at Risk (CVaR), Gini's Mean Difference and minimax measures. In its most basic definition, CVaR is a risk measure of high losses and it has gained considerable popularity in PO literature. CVaR is one of the criteria that we use in this study; and we will discuss it in detail in the following chapter. CVaR-related linear programming models for PO were studied in detail by Mansini et al. (2007). They studied the theoretical properties of models that use multiple CVaR measures as well as their performance on real data. Ogryczak and Ruszczynski (2002) showed that CVaR is consistent with the second degree stochastic dominance and it has attractive computational properties since it can be modeled linearly.

With advances in computing, solving the original mean-variance model has become easier. Steuer et al. (2006) developed a quadratic programming procedure to solve the original model efficiently. Without generating the whole efficient frontier, one can also compute a number of discrete efficient points to represent it. A common approach that is still being widely used for this purpose is the ε-constraint method (Haimes et al., 1971). This method optimizes one of the objectives by systematically changing the levels of the remaining objectives. It is modeled as:

$$\text{Min } f_k(x) \tag{2.4}$$

$$\text{s.t.} \quad f_i(x) \le \varepsilon_i \quad i \ne k \tag{2.5}$$

$$x \in X \tag{2.6}$$

where $f_i(x)$ represents objective function i, ε_i is the value we use to constrain objective i and X represents the set of feasible points. By changing the right-hand side of the constraints, one can achieve points to represent the efficient frontier. We make use of this method to generate solutions for our approaches that produce exact efficient solutions. However, some of the solutions found can be inefficient. To prevent this, we use an augmentation in the objective function to break ties in favor of the objectives treated as constraints. The resulting updated model is:

$$\text{Min } f_k(x) + c \sum_{i \ne k} f_i(x) \tag{2.7}$$

$$\text{s.t.} \quad f_i(x) \le \varepsilon_i \quad i \ne k \tag{2.8}$$

$$x \in X \tag{2.9}$$

where c is a sufficiently small positive constant. This model is referred to as the "augmented ε-constraint method" throughout the thesis.

There can also be variations to ε-constraint method such as in Ballestero and Romero (1996) and Bana e Costa and Soares (2004) where discrete points on the efficient frontier that possess certain characteristics are sought rather than generating a sample from the whole frontier.

Several researchers addressed PO problem with additional constraints. Bertsimas et al. (1999) used a mixed-integer programming approach to construct portfolios that are similar to target portfolios, that have small number of stocks and that require small number of transactions. Bertsimas and Shioda (2009) proposed a method for solving cardinality constrained quadratic optimization problems using a branch and bound implementation. Cesarone et al. (2012) solved the mean-variance portfolio problem with cardinality constraints by a quadratic programming approach.

With time, the use of multiple criteria in PO has emerged. New measures of return and risk were proposed as well as novel criteria that address unconventional issues. The literature on the employment of multiple criteria in finance and PO is covered in Chapter 3. Some researchers used heuristics instead of exact methods to handle multiple criteria and complicating real-life constraints. Approaches such as tabu search, simulated annealing and evolutionary algorithms were used to handle these issues. As an example, Ehrgott et al. (2004) studied PO with five risk and return related objectives. With DM-specific utility functions, they solved it with customized local search, simulated annealing, tabu search and genetic algorithm heuristics. The literature related to our heuristic approach to PO is reviewed in Chapter 4.

As other fields of PO studied by researchers, Ballestro (2001) proposed a stochastic goal programming model to mean-variance PO and utilized the connection between classical expected utility theory and linear weighted goal programming model under uncertainty. Shing and Nagasawa (1999) studied an interactive portfolio selection system based on an α-risk permission maximum return portfolio framework. They used scenarios on security returns and variances with known occurrence probabilities. Dupacova (1999) studied the error and misspecification side of PO problems. Using examples –the Markowitz model, a multi-period bond portfolio management problem and a strategic investment problem-, he presented methods for analysis of results obtained from stochastic programs.

Most of the studies in the literature on PO have traditionally considered a single-period horizon. Recently, SP has come forward as an approach to handle multi-period PO. Constructing scenarios to capture the uncertain progress of economic factors through periods ahead, SP provides us with investment decisions that maximize final expected prosperity. One can refer to Yu et al. (2003) for a survey of SP models in financial optimization. Our literature review related to SP is covered in Chapter 5.

2.3 The Criteria

In this section, we introduce the criteria that we use in several combinations for our studies. Besides expected return, we consider liquidity of the portfolio. We also utilize three risk measures.

Let r_i be the return, $E(r_i)$ be the expected return, and l_i be the liquidity measure of asset *i*, and R_{it} be the return of asset *i* observed in period *t*. Let σ_{ij} be the covariance of the returns of assets *i* and *j*. Let x_i be the proportion of asset *i* in the portfolio where $\sum_{i=1}^{n} x_i = 1$.

2.3.1 Expected Return

Return on the investment is the primary concern of PO. However, this return cannot be known with certainty until after the economical factors that define it are realized. Therefore, expected return comes out as the common basic criterion of most PO problems.

Expected return of a portfolio is given by the weighted expected returns of individual assets by their proportions in the portfolio. It can be expressed as:

$$z_1 = \sum_{i=1}^{n} E(r_i)\, x_i \qquad (2.10)$$

We use percentages to represent expected return values in all of our applications. In all of our studies except for our SP approaches, we estimate expected returns of stocks from historical data. We treat past returns over a certain period as equiprobable scenarios for this purpose. Assuming we are utilizing *k* time periods, the expected return of stock *i* is estimated as:

$$E(r_i) = \frac{\sum_{t=1}^{k} R_{it}}{k} \tag{2.11}$$

In our SP approaches to PO, the expected returns of individual stocks will be derived from the expected return of a market-representative index. We will estimate the expected return of the index by use of a random walk model; and assume that individual stock returns depend on that index with a single index model. The details of this mechanism are discussed in Chapter 5.

2.3.2 Liquidity

Liquidity is the degree a security can be sold without affecting its market price and without loss of value. It can also be defined as the ease a security can be traded within fair price levels. It is particularly important for investors who want to be able to instantly liquidate their assets. Some investors may have frequent payment liabilities, some of which may be without advance notice. Liquidity is usually characterized by the following aspects: time to trade, bid–ask price range (spread) and effect of transaction on price. Sarr and Lybek (2002) reviewed several liquidity measures. Volume-related liquidity measures are generally measured by the volume or quantity of shares traded per time unit. Time-related measures look at the number of transactions or orders per time unit. Spread-related measures study the difference between ask and bid prices with several measurement approaches. There are also multidimensional measures that combine different measures.

We use a liquidity measure that is like a turnover ratio. We use the proportion of shares of a stock that are traded in a fixed time unit among the publicly outstanding number of shares of that stock. We use the most recent month for the number of outstanding shares of a stock. For the number of shares traded in that month, we take the daily average number. We calculate the liquidity measure of each stock using these numbers; the higher the liquidity measure's value is, the more liquid the corresponding stock is. Our liquidity criterion can be expressed as:

$$z_2 = \sum_{i=1}^{n} l_i \, x_i \tag{2.12}$$

2.3.3 Risk Measures

We consider three different risk measures: Variance, CVaR and a measure of under-achievements that we name as Mean Under-Achievement (MUA). We explain each of these measures in detail below:

2.3.3.1 Variance

Variance as a risk measure evaluates how far return values lie from their mean. It is the most traditional and widely-used risk measure of PO. When evaluating the likely return of a portfolio, an investor may want to have an idea about the degree of variation this return will exhibit. When variance is low, expected return values are more reliable. On the other hand, when it is high, the actual return value can deviate from the expected value considerably. Variance of a portfolio is calculated from the covariances between returns of assets contained in the portfolio, and is given by:

$$z_3 = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i \sigma_{ij} x_j \tag{2.13}$$

We compute the covariance terms from historical data. Using *k* past observations, the covariance between returns of assets *i* and *j* is calculated as:

$$\sigma_{ij} = \frac{\sum_{t=1}^{k} [R_{it} - E(r_i)] \left[R_{jt} - E(r_j)\right]}{k-1} \tag{2.14}$$

The resulting covariance matrix *V* is then given by:

$$V = \begin{bmatrix} \sigma_{11} & \sigma_{21} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \cdots & \cdots & & \cdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix} \tag{2.15}$$

2.3.3.2 Conditional Value at Risk

CVaR is a measure of extreme losses. Every investor may want to have an idea about the worst possible position she/he will be in at the end of the investment period. Extreme losses are particularly of interest to investors who consider them as critical. Such investors do not have idle funds; they want to guarantee that they will have certain level of wealth after the investment. CVaR is related to another risk measure, Value at Risk (VaR). CVaR has become the common practice (as opposed to VaR) in the literature due to its desirable mathematical characteristics. VaR is the minimum point α such that, with a preset probability level λ, we will not have losses larger than α. On the other hand, CVaR at confidence level λ is the expected value of losses beyond α. Figure 2 is a normally distributed loss distribution example where VaR and the corresponding CVaR are illustrated.

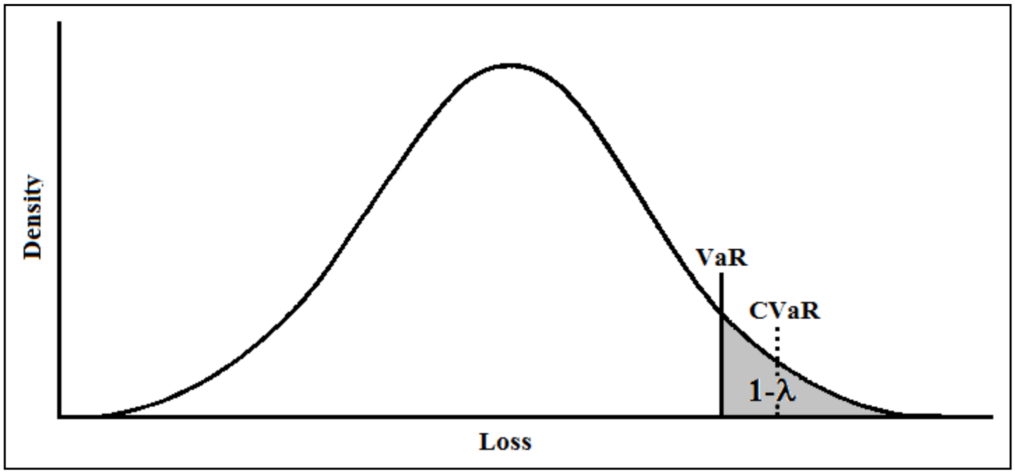


Figure 2. Illustration of VaR and CVaR on an example loss distribution

This definition of CVaR is mostly used in earlier studies on CVaR and can be referred to as the classical definition. Another definition of CVaR at λ probability level, which is used more often in recent studies, is the expected value of losses in the worst ($1-\lambda$) of cases. These two definitions are equal when the random variable under consideration has a continuous distribution. In our studies, we will be referring to the latter definition of CVaR, and using a linear programming model to optimize it. The ease of its computation in the presence of discrete scenarios is another advantage of CVaR over VaR; and we will be making use of this scenario structure in our applications.

Rockafellar and Uryasev (2000) proposed a method to minimize CVaR in a linear model. Here we discuss CVaR and the referred method in the context of PO. See also Rockafellar and Uryasev (2002) for a detailed explanation of CVaR and discussions of its fundamental properties for discrete loss distributions.

Let $f(x,y)$ be the loss corresponding to decision vector x and random vector y. Vector x represents a portfolio that is chosen from a certain subset X of R^n. We can consider x as the vector of proportions of n assets in the portfolio. The results we present below correspond to a given fixed x vector. $y \in R^m$ represents the uncertainties in the market that can affect the loss. If the underlying probability distribution of y is assumed to have density $p(y)$, the probability of $f(x,y)$ not exceeding a threshold τ is given by

$$\psi(x,\tau) = \int_{f(x,y)\leq \tau} p(y)dy \qquad (2.16)$$

Assuming $\psi(x,\tau)$ is everywhere continuous with respect to τ, for a given probability λ, λ-*VaR* and λ-*CVaR* are defined as:

$$\lambda - VaR \; = \min\{\tau \in R : \psi(x,\tau) \geq \lambda\} \tag{2.17}$$

$$\lambda - CVaR \; = (1-\lambda)^{-1} \int_{f(x,y) \, \geq \, \lambda - VaR} f(x,y)\, p(y) dy \tag{2.18}$$

Rockafellar and Uryasev (2000) characterized λ-*VaR* and λ-*CVaR* in terms of a function F_λ as follows:

$$F_\lambda(x,\tau) = \tau + (1-\lambda)^{-1} \int_{y \in R^m} [f(x,y) - \tau]^+ p(y) dy \tag{2.19}$$

where $[\theta]^+$ = maximum $\{0, \theta\}$ (2.20)

It is shown that minimizing this function for a given λ produces the optimal *CVaR* value at the given λ level (Rockafellar and Uryasev, 2000).

For PO, we can take the random vector of uncertainties as r, where $r \in R^n$ represents the returns of n available assets for constructing the portfolio. The loss is then given by the negative of asset returns multiplied with proportions, $-x^T r$. In the presence of q scenarios on asset returns, F_λ can be expressed as:

$$\tilde{F}_\lambda(x,\tau) = \tau + \frac{1}{(1-\lambda)} \sum_{s=1}^{q} [-x^T r_s - \tau]^+ \, p_s \tag{2.21}$$

where r_s is the return vector under scenario s and p_s is the probability of scenario s.

Rockafellar and Uryasev (2000) showed that we can achieve a linear model by using auxiliary variables a_s and converting the above expression to the following objective and constraints:

$$\text{Minimize } \tau + \frac{1}{(1-\lambda)} \sum_{s=1}^{q} a_s \, p_s \tag{2.22}$$

$$\text{s.t. } a_s \geq -x^T r_s - \tau \qquad \forall \, s = 1, \dots, q \tag{2.23}$$

$$a_s \geq 0 \qquad \forall \, s = 1, \dots, q \tag{2.24}$$

We utilize the above CVaR model for our fourth criterion, z_4. In our SP applications, we readily have discrete scenarios to work with. In our other applications where we do not employ scenario generation mechanisms, realizations of returns of assets in past periods are treated as scenarios. We assume that all such scenarios have equal probability. CVaR values presented in the thesis are percentage values.

As evident by now, variance and CVaR differ in that the former evaluates how far the return values lie from their mean and the latter only measures under-achievements. A DM may be interested in both how much variation she/he can expect in the portfolio, and the expectation of the worst cases. Rockafellar and Uryasev (2000) showed that in the case of normal return–loss distribution, variance and CVaR are equivalent in the sense that they result in the same optimal portfolios. However, for other distributions, the DM has sufficient incentive to consider them simultaneously in the PO model. To illustrate, we consider values from two example random loss distributions, one normal and the other chi-squared. Both have expected loss and variance values of 2.013% and 4.057%, respectively. However, at λ = 0.9, the normal distribution has a CVaR value of 5.463% and the chi-squared distribution has a CVaR of 6.685%. We can see that although both distributions have the same expected loss and variance; due to its right-skewness, the chi-squared distribution has a higher CVaR.

2.3.3.3 Mean Under-Achievement

We utilize another linear risk measure of under-achievements. Using a DM-specified minimum-acceptable return level, we find the expectation of under-achievements. We call this measure MUA, and calculate it by utilizing historical returns over k time periods. Accordingly, our fifth criterion is given by:

$$z_5 = \frac{\sum_{t=1}^{k}(R_{it}\, x_i - \pi)^+}{k} \tag{2.25}$$

where π is a return level that the DM sets as her/his critical threshold value.

If we are to compare CVaR and MUA, CVaR does not require the DM to determine a level as the minimum acceptable return. The DM just sets λ as the probability level she/he wants to work with when assessing the risk. Accordingly, we consider CVaR as a more robust measure that can be utilized with various problems of different return/loss distributions. We use MUA in a limited part of our study with genetic algorithm approaches.

In our studies throughout the thesis, we use several different sets of historical data to make the estimations needed to calculate the criteria values. Our basic reason of using different data for different approaches is to make use of more recent data. In addition, some approaches require the use of longer periods of historical data to increase the number of observations used, while others require shorter periods of recent data to obtain estimations that are more representative of the future. We make more specific explanations in the chapters that apply.

CHAPTER 3

EFFECTS OF MULTIPLE CRITERIA ON PORTFOLIO OPTIMIZATION

MCDM has been an important tool in recent studies of finance. Hallerbach and Spronk (2002) discussed why financial decision problems should be handled as multicriteria decision problems. They argued that many financial problems include multiple objectives by nature, and some objectives have different definitions when viewed from different perspectives. They perceive MCDM methods as valuable tools for the whole process of financial decision making. Zopounidis (1999) also pointed out the advantages of MCDM methods in financial management. He discussed multicriteria character of several financial problems. Steuer and Na (2003) provided an extensive bibliography on the application of MCDM techniques to financial issues and problems. As a field in finance, portfolio optimization has also been approached with MCDM tools, especially recently. Steuer et al. (2007a) and (2007b) argued that there are investors who would like to consider several additional criteria besides return such as dividends, liquidity, social responsibility, turnover and amount of shortselling. Numerous researchers confirmed this argument by studying multiple criteria in portfolio optimization. As examples, Roman et al. (2007) used expected return, variance and CVaR as criteria and utilized ε-constraint method where the objective function consists of variance whereas expected return and CVaR are treated as constraints. They applied their approach to stocks from the FTSE index and carried out in and out-sample analysis on their results. Xidonas et al. (2011) considered expected return, MAD, dividend yield and market risk in the presence of several constraints on portfolios. Binary variables included in some constraints led to a mixed-integer multiobjective linear programming model which was solved with augmented ε-constraint method. An interactive filtering procedure was applied on the generated solutions to guide the DM to her/his most preferred solution. Xidonas et al. (2010) proposed a similar approach with two additional criteria: relative price/earnings ratio and marketability. Fang et al. (2009) also employed ε-constraint method to generate efficient solutions with three criteria: net expected return after transaction costs are deducted, semi-absolute deviation and a fuzzy liquidity measure of turnover rates of securities.

Despite the increasing use of MCDM in PO, to the best of our knowledge, there are no studies that explicitly discuss the effects of using multiple criteria on the portfolio of an investor. Making use of single-period PO settings, we study the effects of different and additional criteria on objective and decision spaces. Different combinations of expected return, variance, liquidity and CVaR are used as test problems. We also consider the behavior of the criteria in the presence of cardinality and weight constraints, and provide comparative discussions.

We make tests with stocks from ISE and use the augmented ε-constraint method to find efficient solutions. In all of our applications with two, three and four criteria, we use the same set of stocks from ISE as available assets. We choose 70 stocks from ISE (see Table 17 in Appendix A), and use their monthly percentage return data between January 2005 and December 2010[1] to estimate their expected returns and the covariance matrix. The returns of the stocks during these 72 months are treated as equiprobable scenarios for CVaR and 90% probability is used. For the liquidity measure, we use the daily average number of shares traded[1] and the total number of outstanding shares[2] corresponding to June 2011. For our tests, we use different combinations of criteria in two, three and four-criteria settings.

The work covered in this section draws on our paper Tuncer Şakar and Köksalan (2013a).

3.1 Two-Criteria Models

In this section, we show the effects of using different bicriteria combinations of expected return, variance, liquidity and CVaR measures. For all cases, we study the conflict between the criteria. Conflict in our context refers to the inverse relation between criteria considered in a problem. Better values of a criterion cannot be achieved unless sacrifices are made from the other(s). If this was not the case, we could improve all criteria simultaneously and attain a single optimal solution. However, in reality, we can attain a set of efficient solutions in the presence of multiple criteria. As we move along this set of solutions, some criteria values will improve and others will worsen. The degree of conflict between two criteria can be of varying degrees. A strong conflict corresponds to the case where the two criteria force each other to very poor values in order to attain good values themselves. A weak conflict, on the other hand, shows itself in a more limited range of criteria values.

3.1.1 Expected Return–Variance

We first find efficient expected return–variance solutions with our stocks from ISE. We treat minimizing variance as the objective function and lower bounds on expected return as constraints in the augmented ε-constraint method. We generate 20 efficient solutions that are evenly spaced in the expected return range. The model used for this purpose is provided as a sample augmented ε-constraint method used in our applications:

$$\text{Min} \sum_{i=1}^{n}\sum_{j=1}^{n} x_i\,\sigma_{ij}\,x_j + c\sum_{i=1}^{n} E(r_i)\,x_i \tag{3.1}$$

$$\text{s. t.} \sum_{i=1}^{n} E(r_i)\,x_i \geq \varepsilon \tag{3.2}$$

[1] http://borsaistanbul.com/en/data/data/equity-market-data/equity-based-data

[2] http://www.mkk.com.tr/wps/portal/MKKEN/InvestorServices/eDATACapitalMarketsDataBank

$$\sum_{i=1}^{n} x_i = 1 \tag{3.3}$$

The variables and parameters of the above model were defined previously in Chapter 2.

The minimum-variance portfolio consists of 15 stocks and the maximum-expected return portfolio consists of one stock in our case. We observe a decline in the number of stocks as we move towards the maximum-return portfolio. We also want to see how the two criteria will behave in the presence of constraints on the maximum number of stocks (referred to as cardinality constraints) and the maximum weight a stock can take (referred to as weight constraints). We expect the weight constraints to prevent the best possible values of expected return since it will not be possible to fully invest in the stock with the maximum expected return. The best values of variance will also be expected to worsen, this time because of cardinality constraints that prevent full differentiation. The average number of stocks in the 20 efficient solutions generated is approximately six, so we choose six as the maximum number of stocks allowed in a portfolio since this will be a reasonably restrictive constraint. And for the weight constraint, we choose 0.4. This weight is small enough to disturb the portfolios that are close to the maximum-expected return portfolio and large enough to give flexibility to portfolios that are also constrained by the cardinality constraint. We again generate 20 efficient solutions that are equally spaced in the expected return range by using a mixed integer augmented ε-constraint model to handle the cardinality constraint. The model is given by (3.1)-(3.3) extended by the following constraints:

$$x_i \leq 0.4\, b_i \qquad \forall\, i = 1, \ldots, n \tag{3.4}$$

$$\sum_{i=1}^{n} b_i \leq 6 \tag{3.5}$$

$$b_i \in \{0,1\} \qquad \forall\, i = 1, \ldots, n \tag{3.6}$$

Figure 3 shows the efficient frontiers with and without the constraints. We see that the two efficient frontiers converge fairly well in the range of the constrained case. Concerning expected return, we see that the constraints make a substantial portion of the high expected return values unattainable. On the other hand, the best variance values are very close with and without constraints. This shows that in this case, a maximum of six stocks are sufficient to decrease variance near its best value. Another observation is that, with constraints, the worst variance is substantially lower than the case of no constraints (variance reaches very high values without constraints). This is due to the fact that the corresponding high expected returns are now unattainable. Since we have to invest in at least three stocks because of the weight constraint, variance no longer assumes very high values that result from portfolios having the highest expected return values by investing in one or two stocks. For this case, in general we can say that expected return and variance are highly conflicting, and cardinality and weight constraints decrease the degree of this conflict.

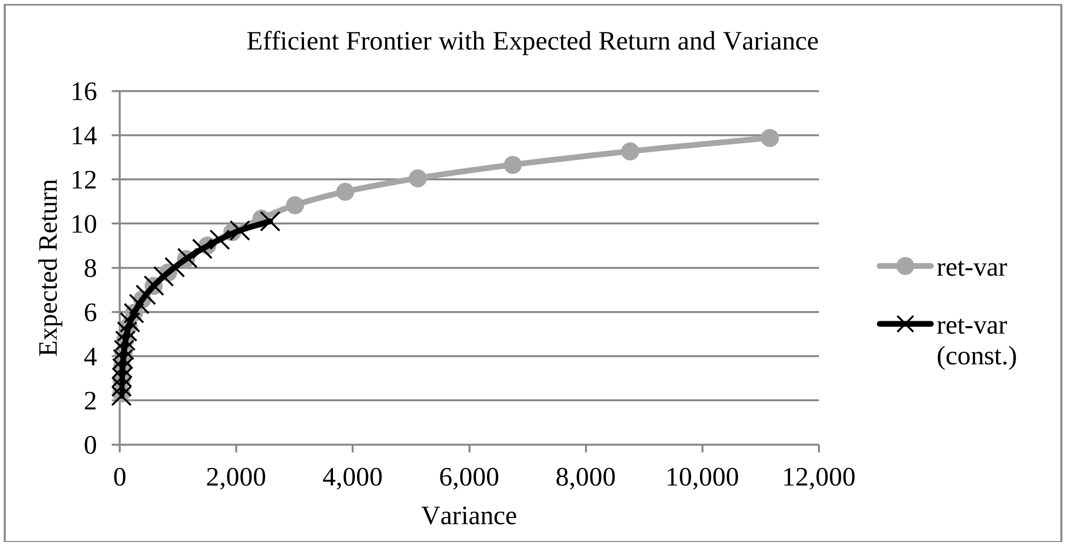


Figure 3. Efficient frontier with expected return (ret) and variance (var) with and without constraints

We also want to compare the compositions of portfolios with different expected return/variance levels. We expect portfolios having high expected return to negatively correlate with portfolios with low variance in terms of their asset contents. We use the correlation coefficient (*r*) to measure the correlation between the weights of stocks in pairs of portfolios. For both unconstrained and constrained cases, we have 20 efficient solutions. For each case, we sort our solutions in increasing order of expected return (and variance) and denote them as s_1, s_2, …, s_{20}. We compare solution s_1 with s_{20}, s_2 with s_{19} and so on, and obtain ten pairs for both cases. The first pairs correspond to the portfolios with the farthest expected return values and the last pairs correspond to those with the closest expected return values.

After calculating *r* values between these pairs of portfolios, we also check for their significance. We test for the statistical significance of r using the test statistic $\frac{\hat{r}}{\sqrt{\frac{1-\hat{r}^2}{n-2}}}$ which approximately has *t* distribution with *n*–2 degrees of freedom under the null hypothesis of zero correlation (see Kendall and Stuart, 1979, p. 501–503). In our case, since we have 70 stocks to choose from, *n*=70. Our hypotheses are:

$H_0 : r = 0$
$H_A : r < 0$

When we look at the correlation coefficients, we see that for both unconstrained and constrained cases, there is negative correlation for pairs with farthest expected return values as expected. We also see that the correlation increases as the expected returns (and variances) approach each other and reach very high levels for the last pairs. However, the results corresponding to negative correlations are not statistically significant. Studying the compositions of portfolios, we realize that there are a number of stocks that are never chosen in any portfolio. These stocks become irrelevant to our problem; if we exclude them before the optimization process, we achieve exactly the same results. The fact that these stocks have zero weight in both portfolios of a pair increases the correlation between them. If we keep adding other similar irrelevant assets from the market to our asset pool, we can further increase correlations. Therefore, we also study the case of excluding irrelevant stocks and considering the remaining 21 for the unconstrained case and 14 for the constrained case. The progress of correlations from negative values to high positive values is still observed. For both cases, excluding irrelevant stocks results in decreased correlations as expected. Another observation is that the correlations in the constrained case are lower than the unconstrained case. However, negative correlations are still insignificant.

3.1.2 Expected Return–Liquidity

Financial markets are believed to demand extra return from illiquid stocks for their inconvenience in trading, so we expect to see conflicting behavior between these two criteria. Using liquidity in the objective function and expected return in the constraint in the augmented ε-constraint method, we generate 20 efficient solutions that are evenly spaced in the expected return range. Both criteria assume their best values when a single stock with the best expected return or liquidity is selected; so these stocks constitute the end point portfolios of the efficient frontier. All other portfolios consist of these two stocks in varying proportions. For the case with constraints, our previous weight constraint is appropriate since it requires at least three stocks in a portfolio. The cardinality constraint of maximum six stocks is also applied even though we do not expect it to be binding. Weight constraints are expected to have similar effects on liquidity as expected return; maximum liquidity values will no longer be possible.

Figure 4 illustrates the efficient frontiers with and without constraints. We see that expected return and liquidity conflict as expected. Without constraints, the efficient frontier is approximately linear, whereas constraints disturb this property. Efficient frontiers with and without constraints converge well only for medium values of expected return and liquidity. We observe that the ranges of both criteria narrow down from both sides in the presence of constraints. The decrease in the upper bounds of the two criteria is due to weight constraints as previously explained. The other consequence of this situation is that, since we cannot obtain the best expected return (liquidity) values, liquidity (expected return) values do not need to worsen as much as the unconstrained case.

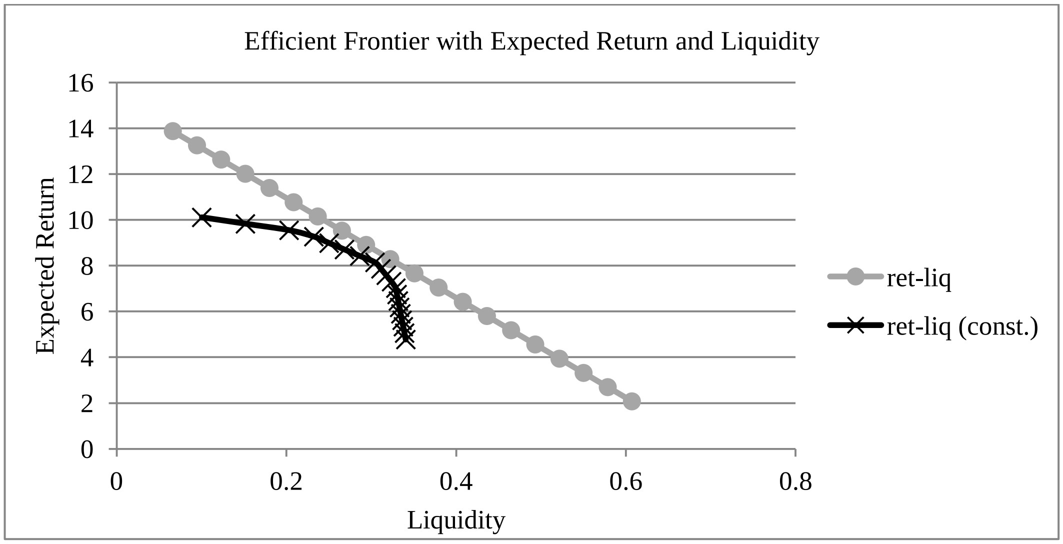


Figure 4. Efficient frontier with expected return (ret) and liquidity (liq) with and without constraints

3.1.3 Expected Return–CVaR

In this case, CVaR is used as the risk measure along with the traditional expected return criterion. As stated before, 90% probability is used for CVaR. We generate 20 efficient solutions that are evenly spaced in the expected return range by using CVaR in the objective function and expected return in the constraint in the augmented ε-constraint method. The average number of stocks of these 20 portfolios is also realized as six, and 0.4 is observed to be an appropriate maximum weight constraint. The use of the same constraints will be preferred for all cases for purposes of consistency, and it will also provide us with two-criteria solutions compatible to be used together to generate three and four-criteria efficient solutions. Hence, constraints in all cases of this study correspond to the same cardinality and weight constraints. Figure 5 illustrates the expected return–CVaR efficient frontiers with and without constraints.

Looking at Figures 3, 4 and 5, we can see that the range of expected return without constraints is approximately the same in all cases. The response of the expected return–CVaR case to constraints resembles the expected return–variance case. We see that although constraints prevent the best values of expected return, we can still attain near-best CVaR values. In addition, when we add the constraints, the worst CVaR value is lower than the case of no constraints. This can again be attributed to the fact that, since the weight constraint prevents expected return to reach its highest values, CVaR no longer needs to assume very high values. Constraints decrease the degree of conflict between expected return and CVaR.

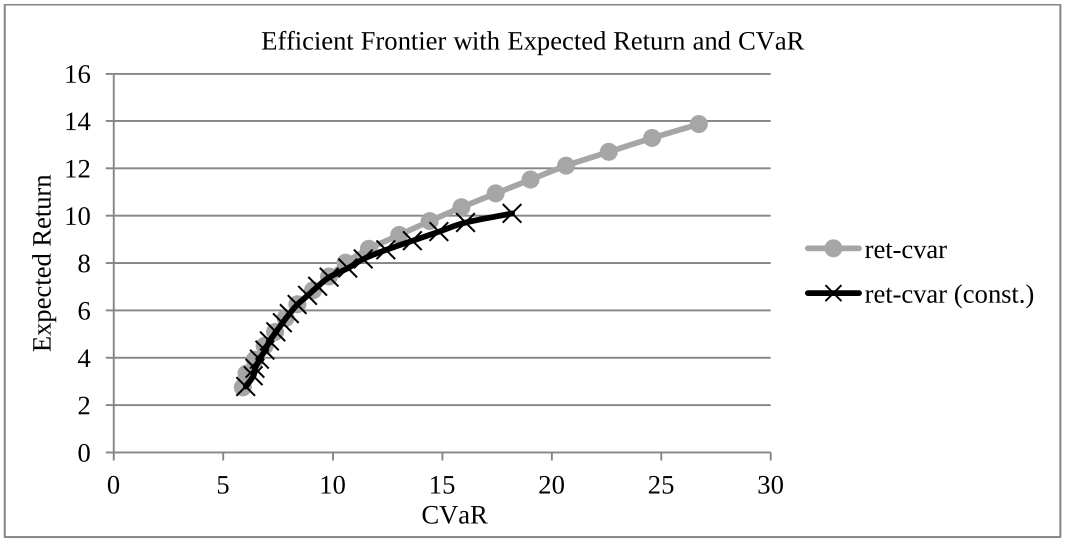


Figure 5. Efficient frontier with expected return (ret) and CVaR (cvar) with and without constraints

3.1.4 Variance–Liquidity

Using variance in the objective function and liquidity in the constraint in the augmented ε-constraint method, we generate 20 efficient solutions that are evenly spaced in liquidity. Figure 6 shows the efficient frontiers with and without the constraints. We see that the two criteria exhibit conflicting behavior in both cases. A possible explanation for this behavior may be that, illiquid stocks are expected to offer investors additional stability. An illiquid stock with high risk would not be considered an attractive investment. In Figure 6, firstly we observe that the range of variance without constraints is quite narrow; suggesting that variance and liquidity's conflict is not very strong in this case. Also, we see that the weight constraint reduces the best possible liquidity value greatly whereas the minimum variance values with and without constraints are very close. The two efficient frontiers converge well for low liquidity values, but variance deteriorates considerably after a point. Constraints increase the degree of conflict between variance and liquidity in this case. We have to sacrifice great amounts in variance to reach the best values of liquidity in the presence of constraints.

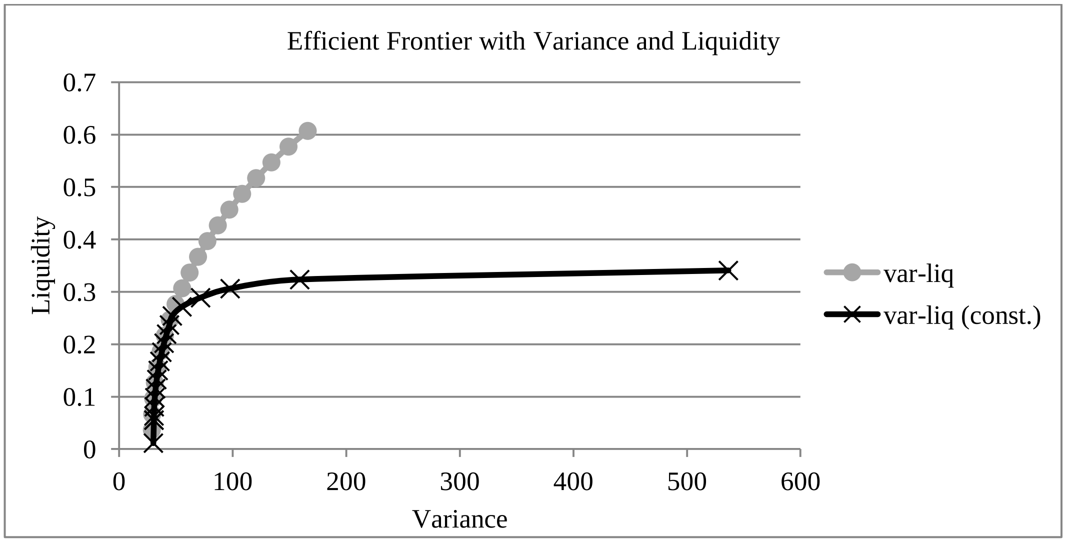
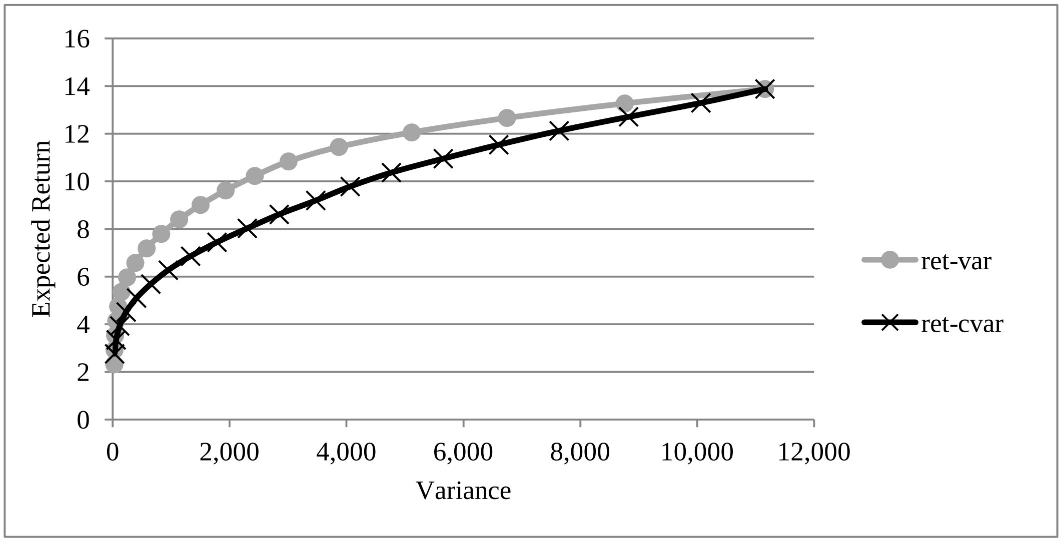


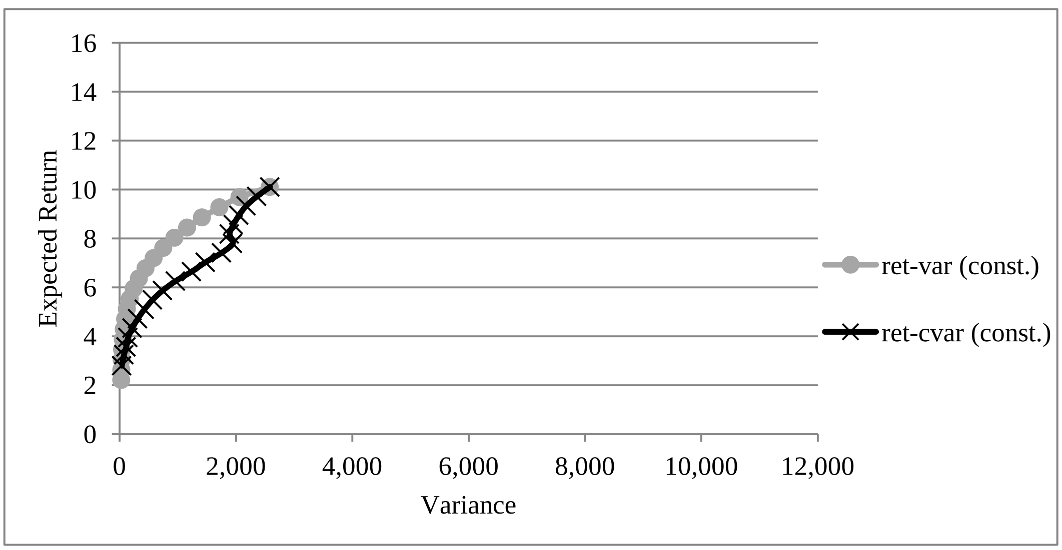
Figure 6. Efficient frontier with variance (var) and liquidity (liq) with and without constraints

3.1.5 Variance–CVaR

Before using variance and CVaR together, we first make computational studies with our data to see if the two risk measures will result in similar optimal portfolios. We compare the efficient portfolios of the expected return–variance and expected return–CVaR cases. If the two risk measures lead to similar efficient portfolios, the CVaR (variance) values of efficient expected return–variance (expected return–CVaR) portfolios should be comparable to those values of efficient expected return–CVaR (expected return–variance) portfolios. Figures 7 and 8 help us to make those comparisons. In Figure 7, we plot efficient expected return–variance and expected return–CvaR solutions in expected return and variance axes, with and without constraints. In Figure 8, we plot the same solutions in expected return and CVaR axes. As Figure 7 illustrates, efficient expected return–CVaR portfolios cannot approach the variance levels of efficient expected return–variance portfolios, and constraints increase the divergence further. Likewise, efficient expected return–variance portfolios cannot converge the CVaR levels of efficient expected return–CVaR portfolios as evident from Figure 8. Constraints again increase the divergence. In both figures, the cases with constraints are drawn in the same scale as the unconstrained cases to illustrate the effects of constraints on the range of criteria as well. Results show that our data taken from ISE for the specified period does not demonstrate normal return distribution. This result makes variance and CVaR suitable risk measures to be used together.

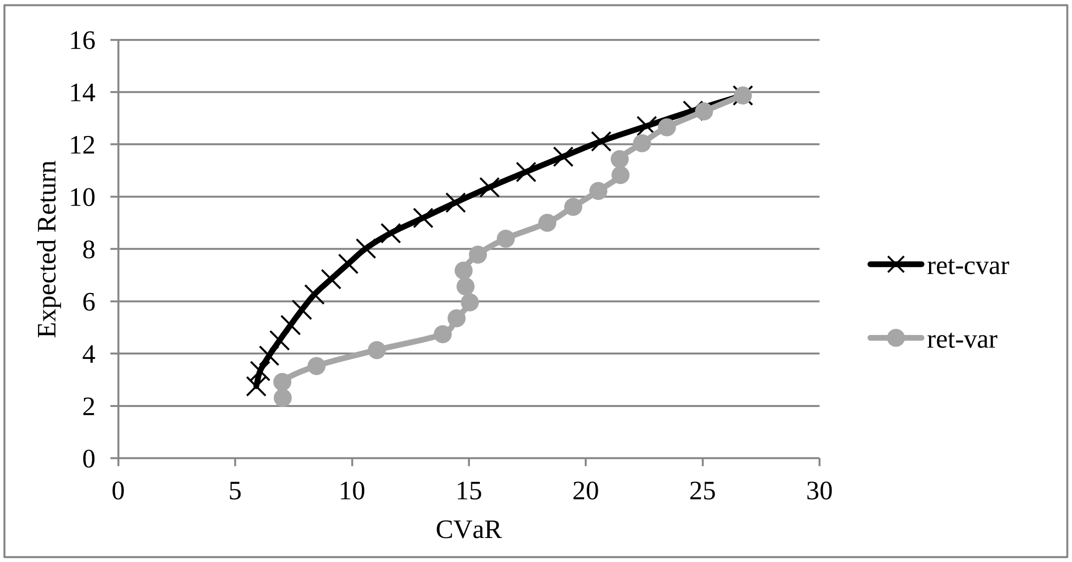


(a)

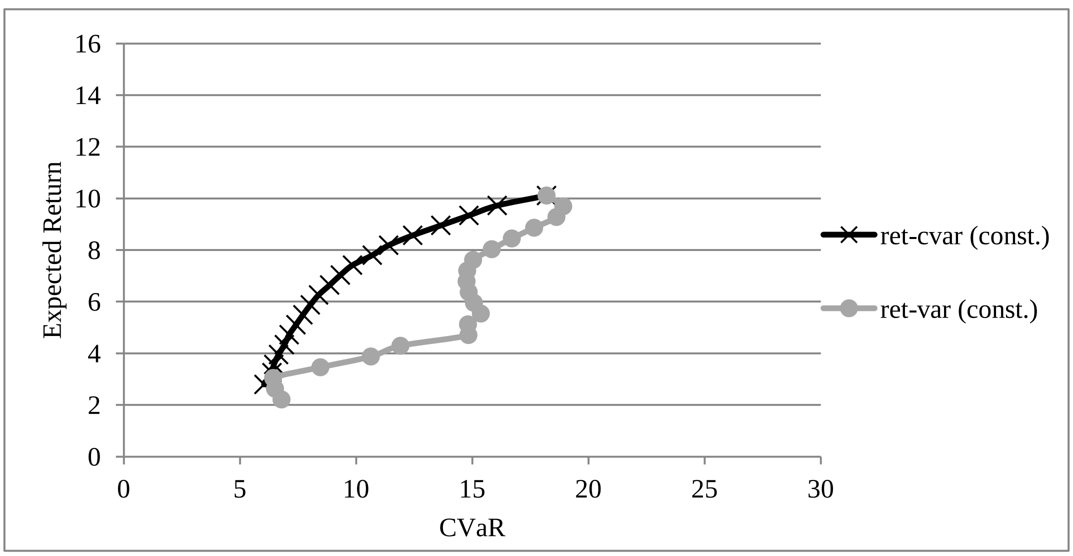


(b)

Figure 7. Comparing expected return–variance (ret–var) and expected return–CVaR (ret–cvar) efficient portfolios in expected return and variance without (a) and with (b) constraints



(a)



(b)

Figure 8. Comparing expected return–variance (ret–var) and expected return–CVaR (ret–cvar) efficient portfolios in expected return and CVaR without (a) and with (b) constraints

We also want to compare the compositions of portfolios with variance and CVaR; we expect to see them result in different portfolios. We will again use *r* for this purpose. This time, we will form our pairs of portfolios by choosing expected return–variance and expected return–CVaR efficient portfolios with the same expected return value. For both unconstrained and constrained cases, we use expected return values that are 0.5 points apart and obtain 22 such values with the unconstrained case and 15 values with the constrained case. We sort them in increasing order of return. Our hypotheses are:

$H_0 : r = 0$
$H_A : r > 0$

For both cases, we observe high correlation for low and high return values and low correlation for intermediate values. For two out of 22 pairs of portfolios with no constraints, we fail to reject zero correlation; and with constraints, this number is two out of 15. Although we reject our null hypothesis for most of the pairs, we still cannot conclude that there is high correlation. It is a very strong argument to expect variance and CVaR to result in totally uncorrelated portfolios. The magnitudes of our *r* measure can be more useful here to give an idea about the strength of correlation, and we observe low *r* values except for portfolios with low and high return.

Following our previous line of thought, we can again exclude the stocks that are not chosen in any efficient portfolio. We have 20 stocks remaining in the unconstrained case and 15 in the constrained case. For the unconstrained and constrained cases, the number of pairs that we fail to reject zero correlation increases to seven and nine, respectively. So in our case, constraints increase the difference in portfolio compositions with variance and CVaR. Overall, we conclude that the compositions of portfolios show that variance and CVaR lead to reasonably different portfolios and they can be used in the same model.

To construct a discrete representation of the variance–CVaR efficient frontier, we solve for minimum variance for 20 evenly spaced CVaR values. We apply the same procedure in the presence of constraints as well. Figure 9 shows the efficient frontiers with and without constraints. We see that the two risk measures exhibit conflicting behavior. However, the range of both criteria is observed to be very narrow in contrast to the previous cases that included them. The range of variance in the expected return–variance model was [28.52–11,160.40] and the range of CVaR in the expected return–CVaR case was [5.90–26.72]. So, variance and CVaR cover each other to a considerable extent and conflict in a limited region for both criteria. Constraints restrict this region ever further for CVaR whereas the range for variance shifts to higher values.

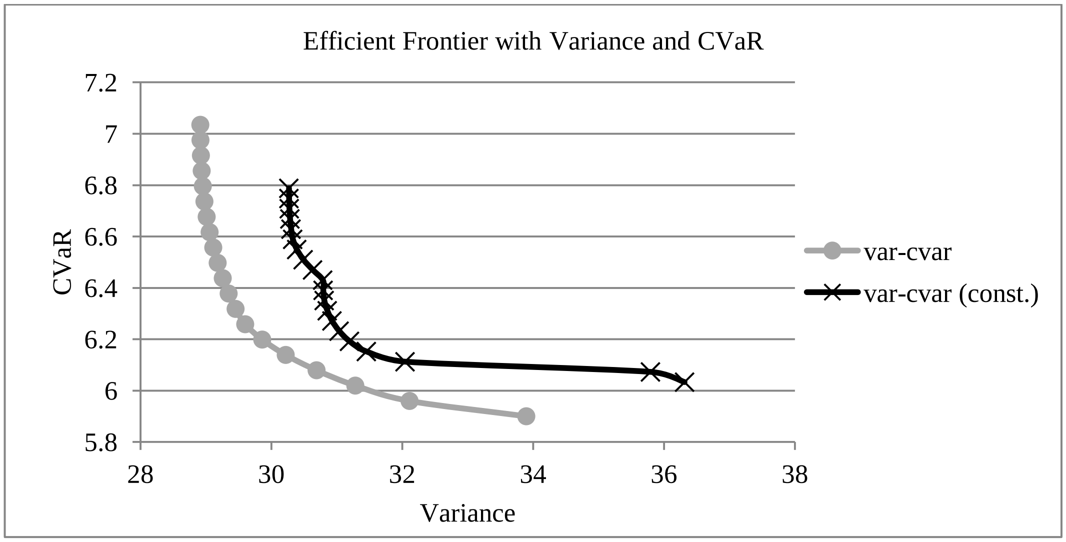


Figure 9. Efficient frontier with variance (var) and CVaR (cvar) with and without constraints

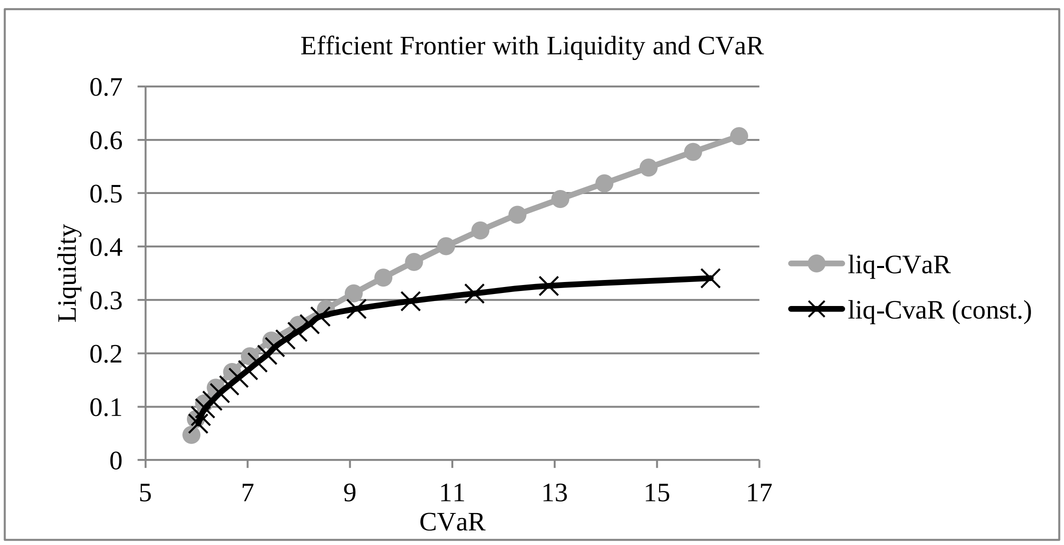


Figure 10. Efficient frontier with liquidity (liq) and CVaR (cvar) with and without constraints

3.1.6 Liquidity–CVaR

We generate 20 efficient solutions that are evenly spaced in liquidity using CVaR in the objective function and liquidity in the constraint in the augmented ε-constraint method. Figure 10 shows the efficient frontiers with and without the constraints; the criteria conflict properly to allow their use together in the same model. We can observe that, although the weight constraints reduces the best possible liquidity value greatly, the best CVaR value with constraints remains very close to the unconstrained case. The range of CVaR with constraints is also similar to the unconstrained case. Another observation is that the two efficient frontiers converge well for low liquidity and CVaR values, but after a point, constraints lead to increasingly lower liquidity values for given CVaR levels.

3.2 Three-Criteria Models

In this section, we look at two representative three-criteria models. In addition to our two traditional criteria, expected return and variance, we first consider liquidity and then CVaR as the third criterion. We saw in Section 3.1 that all criteria considered conflict pairwise; and thus are compatible to be used together in the same model. To find efficient solutions with three criteria, we employ the augmented ε-constraint method with a two-step procedure. In the first step, for every efficient point of the two-criteria models, we use the values of the two criteria as constraints and solve for the optimum value of the remaining criterion. We treat the worst value of the optimized criterion obtained from these solutions as the nadir point in that criterion. This would have corresponded to the exact nadir point had the problem been discrete (see Ehrgott and Tenfelde-Podehl, 2001). The quality of the obtained nadir value can be made as precise as desired by controlling the amount of discretization. As a result, this procedure also ensures that we cover the efficient range of the three criteria as precisely as desired. The first step may still leave some unrepresented efficient regions between the nadir and the ideal values of the criteria. Therefore, in the second step we find additional efficient solutions to represent the possible under-represented regions of the criteria by solving models imposing bounds in the desired regions of the criteria.

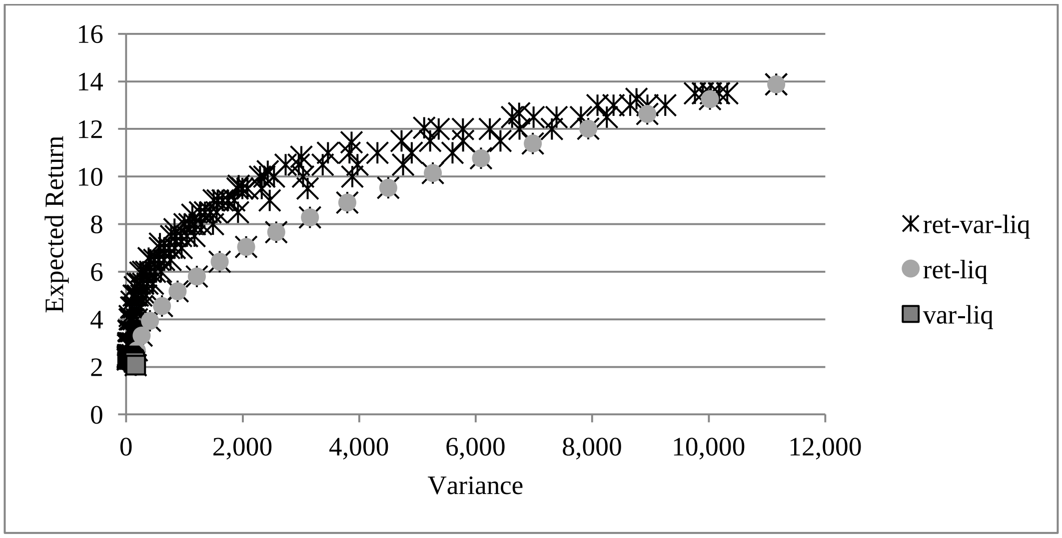
In both three-criteria models, we first study if considering a third criterion brings improvements in its values compared to the two-criteria case. Without optimizing it, we find the corresponding values of the third criterion in the two-criteria efficient solutions. Then, we compare these values to its optimum values found given the same levels of the other two criteria. As the second analysis, we compare the ranges of criteria in two and three-criteria cases. We make both of our analyses in the presence of constraints as well.

3.2.1 Expected Return–Variance–Liquidity

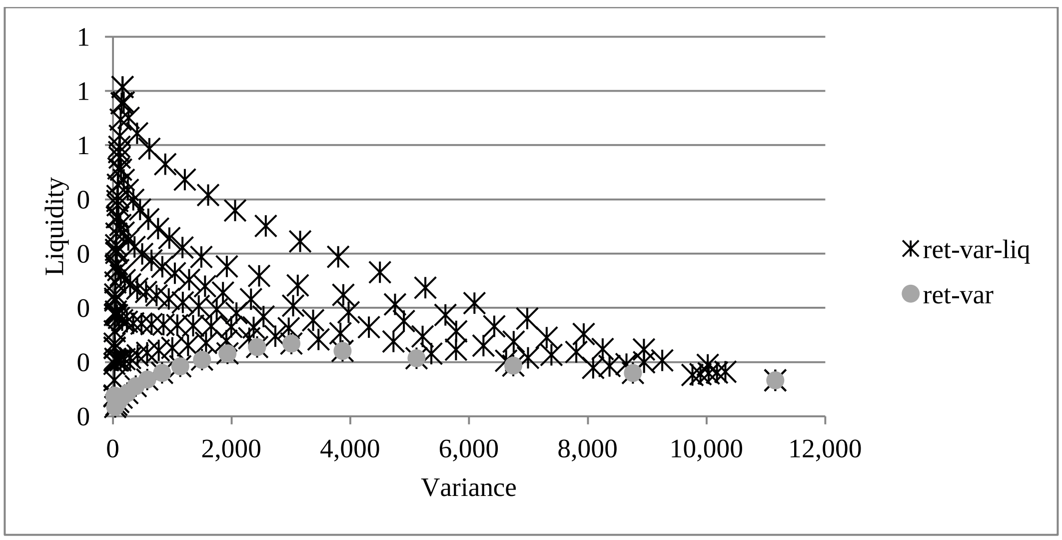
From Section 3.1, we have 20 efficient solutions for expected return–variance, expected return–liquidity and variance–liquidity models, both with and without constraints. First with no constraints, we find the values of the missing third criterion in these cases and compare these to its optimal values found given the levels of the two other criteria. We do not see substantial differences; the three-criteria model fails to bring big improvements in the third criterion within the efficient frontier of the original criteria. For expected return, the average improvement we achieve is 0.1172%; for variance the average improvement is 0.0007% and for liquidity it is 0.1961%. This reveals that for given values of two criteria, there do not exist alternative solutions that are considerably better than others in terms of the third criterion. With constraints, the average improvements in expected return, variance and liquidity are 0.0421, 0.0016 and 0.0540, respectively. We see that the improvement in variance increases with constraints whereas it decreases for expected return and liquidity, but the improvement numbers are very small in both cases to make a meaningful interpretation. We conclude that the changes in improvements are unpredictable and result from how the criteria in two-criteria models behave in the presence of constraints.

Interpreting this observation as there is no need to include the third criterion in the optimization process would be misleading. If we do not consider the third criterion, we cannot cover the range it will assume in the three-criteria case. The best possible values of it will be overlooked; and these may be of interest to the investor. Figure 11 illustrates this situation with our data. We show our results in two-dimensional graphs for easier visualization purposes. In Figure 11(a), when we compare the return ranges of variance–liquidity and expected return–variance–liquidity models, we see that expected return values are realized at very low levels in the former model. Also, expected return–variance–liquidity model can reach lower levels of variance than the return–liquidity model. The lowest variance value the return–liquidity model can obtain is 166.16 while its lowest value with the three criteria model is 28.92. Figure 11(b) shows the narrow range liquidity assumes in the expected return–variance model as opposed to the three-criteria case.

In three dimensions, Figure 12 shows the efficient points achieved by the three-criteria model together with the efficient points achieved by the three two-criteria models. We can see that two-criteria models can only cover small ranges.



(a)



(b)

Figure 11. Comparing the (a) return (ret), variance (var) and (b) liquidity (liq) ranges of two-criteria models with the three-criteria model

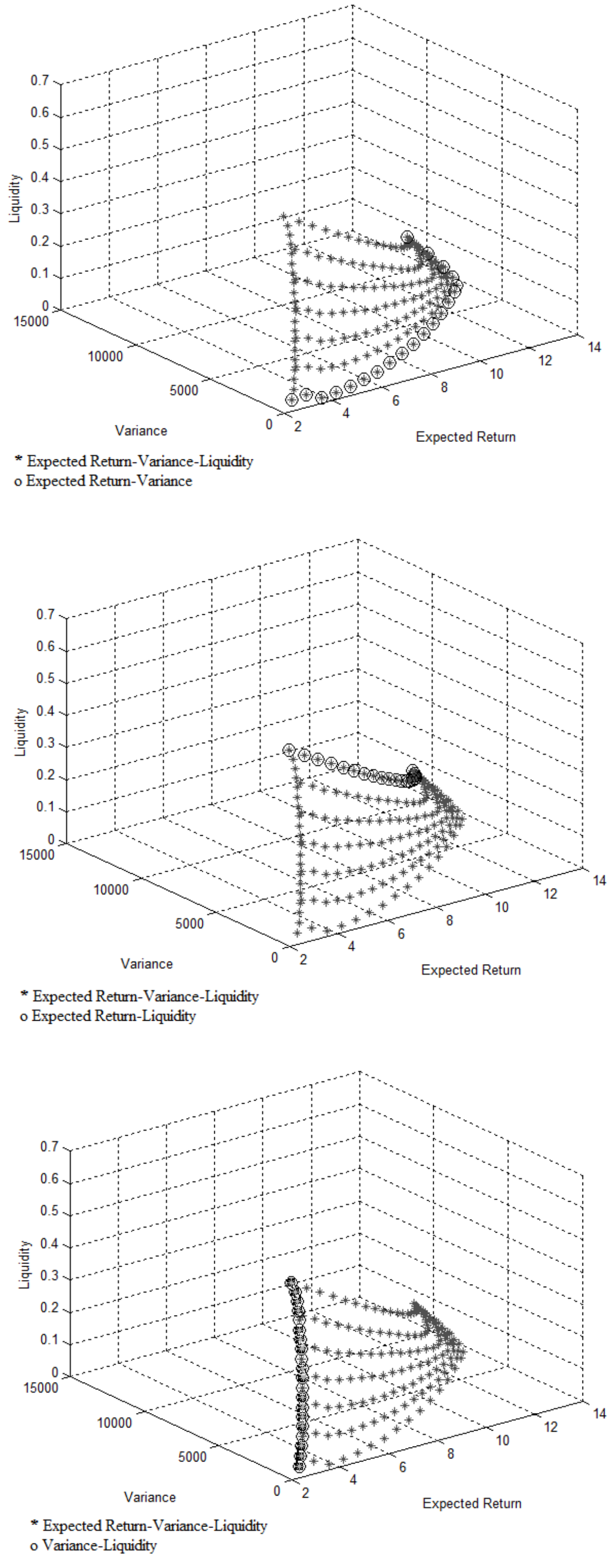


Figure 12. Efficient points of expected return–variance–liquidity model compared with expected return–variance, expected return–liquidity and variance–liquidity models

Table 1(a) has ranges of the criteria in solutions of the three-criteria model and also the ranges of the missing third criterion in two-criteria models. We can see that expected return is the criterion that the two-criteria model fails to cover most with approximately only a 3% coverage. For liquidity, the approximate coverage percentage is 20% and for variance it is 99%. Although the coverage for variance seems very high, it is partly due to its extremely high upper bound caused by the maximum return portfolio in the expected return–variance case. There is considerable difference between the best variance values of the three-criteria and the expected return–liquidity models. To see the effects of constraints on our findings, we also report the ranges and the coverage percentages in the presence of constraints in part (b) of Table 1. We again see that two-criteria models miss regions that are good in the omitted criterion. In our case, constraints increase the expected return coverage of the variance–liquidity model and the liquidity coverage of the expected return–variance model. However, there is a decrease in the variance coverage of the expected return–liquidity model.

Analysis of the three-criteria models' ranges in Table 1 also gives us information about the effects of constraints with criteria of expected return, variance and liquidity. We see that the lower bound of the range of expected return increases to a better value with constraints, but the upper bound worsens. The lower bound for variance increases (becomes a worse value), but its upper bound decreases to a better value. For liquidity, both the lower and the upper bounds worsen.

Table 1. Ranges of criteria realized with two and three-criteria models (ret–var–liq) and the % coverage achieved by two-criteria models

(a) without constraints

	Expected Return	Variance	Liquidity
Three-criteria model range	2.08–13.88	28.92–11,160.30	0.02–0.61
Expected return–variance model range (% coverage)			0.02–0.13 (19.84%)
Expected return–liquidity model range (% coverage)		166.16–11,160.40 (98.77%)	
Variance–liquidity model range (% coverage)	2.08–2.46 (3.29%)		

(b) with constraints

	Expected Return	Variance	Liquidity
Three-criteria model range	2.21–10.11	30.27–2,582.76	0.01–0.34
Expected return–variance model range (% coverage)			0.01–0.10 (27.01%)
Expected return–liquidity model range (% coverage)		537.82–2,582.77 (80.12%)	
Variance–liquidity model range (% coverage)	2.21–4.77 (32.43%)		

3.2.2 Expected Return–Variance–CVaR

In this model we use two risk measures besides our expected return criterion. Upon finding our efficient points, we first check for the improvement they bring in the formerly left-out criterion. Similar to the previous three-criteria case, we fail to see significant values. For the points we consider, the average improvement is 0.0563% in expected return, 0.0074% in variance and 0.0910% in CVaR. When we look at the effects of constraints on these observations, we see that improvements in expected return, variance and CVaR become 0.0153%, 0.0059% and 0.0478%, respectively. We see that the improvements in all criteria decrease in the presence of constraints. However, this decrease is very slight and not generalizable. We conclude that with constraints too, for given values of two of the expected return, variance and CVaR criteria, there are no alternative solutions that can improve the third criterion greatly.

Despite this conclusion, in this three-criteria case too, not considering the third criterion will result in overlooked regions in its range. Table 2 illustrates the ranges of criteria in solutions of the three-criteria model and also the ranges of the missing third criterion in two-criteria models, with and without constraints. We see that the range of expected return in the three-criteria model is again missed to a great extent by the two-criteria model. In the presence of constraints, the variance–CVaR model can cover a wider range of expected return, but the coverage is still below 8%. Another observation is that, both with and without constraints, the ranges of both risk measures are covered well with two-criteria models although best values of both are missed to some extent. This is consistent with our discussions in Section 3.1.5 where we saw that variance and CVaR conflict in a very limited region and they cover each other well. An implication of this result can be as follows: If the DM is not interested in the best possible values of variance (CVaR), solutions of the two-criteria model expected return–CVaR (expected return–variance) can be sufficient for her/him to base her/his decisions on.

Table 2. Ranges of criteria realized with two and three-criteria models (ret–var–cvar) and the % coverage achieved by two-criteria models

(a) without constraints

	Expected Return	Variance	CvaR
Three-criteria model range	2.32–13.88	28.92–11,160.30	5.90–26.72
Expected return–variance model range (% coverage)			7.02–26.72 (94.65%)
Expected return–CVaR model range (% coverage)		33.90–11,160.40 (99.96%)	
Variance–CVaR model range (% coverage)	2.31–2.58 (3.82%)		

(b) with constraints

	Expected Return	Variance	CvaR
Three-criteria model range	2.21–10.11	30.27–2,582.71	6.03–18.91
Expected return–variance model range (% coverage)			6.43–18.91 (96.96%)
Expected return–CVaR model range (% coverage)		36.32–2,5282.71 (99.76%)	
Variance–CVaR model range (% coverage)	2.21–2.83 (7.81%)		

Similar to Table 1 of Section 3.2.1, analysis of the three-criteria models' ranges in Table 2 also gives us information about the effects of constraints with criteria of expected return, variance and CVaR. Here we also support this analysis by the help of Figure 13, where we can see three-criteria solutions with and without constraints in three dimensions. The lower and upper bounds of the range of expected return decreases to lower values with constraints. The lower bound for variance increases (it becomes a worse value), but its upper bound decreases to a better value. For CVaR, the lower bound increases whereas the upper bound decreases with constraints. The decrease in the upper bound of variance and CVaR are most easily seen on Figure 13.

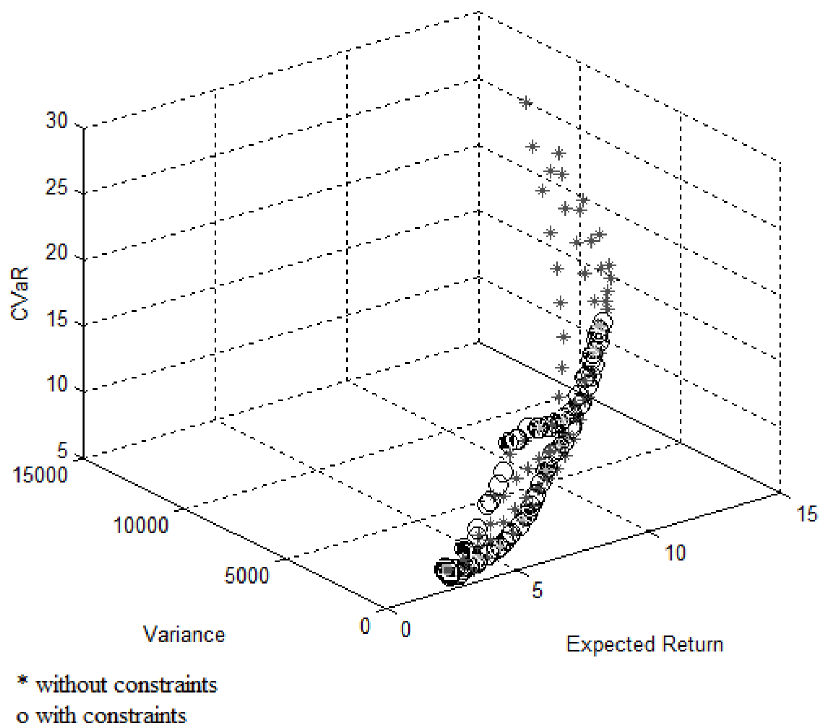


Figure 13. Efficient points of expected return–variance–CVaR model with and without constraints

3.3 Four-Criteria Model

The four-criteria case can also be solved with the augmented ε-constraint method, though with increased computational burden. Using the values of the criteria in efficient three-criteria solutions as constraints, we can solve for the optimum value of the fourth criterion. Then we can solve for intermediate values of criteria to represent all regions of their ranges. We expect to observe results similar to the three-criteria cases. We do not expect substantial improvements in the fourth criterion values corresponding to the efficient solutions of the three criteria cases; but we expect an enlarged efficient region. Looking at some of our results, solving for maximum liquidity given the expected return, variance and CVaR values of the three-criteria model can only bring an average of 0.4208% improvement in liquidity values. But the expected return–variance–CVaR model realizes highest liquidity at 0.13 whereas the true maximum in the four-criteria case is 0.61. Likewise, although CVaR values of expected return–variance–liquidity case only improve by 0.0441% on the average, CVaR can achieve its minimum at 5.90 instead of 7.02 of the three-criteria model. Other criteria exhibit similar behavior too. When we look at the effects of constraints on these observations, we see that they decrease the improvement to 0.0662% for liquidity and 0.0190% for CVaR. But similar to the cases with no constraints, there are substantial differences between the best values attained by three and four-criteria models: 0.10 vs. 0.34 for liquidity and 6.42 vs. 6.03 for CVaR.

We conclude that, for stocks traded on ISE, the four criteria we have considered in this study are compatible for use in the same model if they are of concern to the DM. The inclusion of the fourth criterion to three-criteria cases does not result in discovering alternative solutions that are substantially better in the fourth criterion, but it will result in additional regions. These conclusions are also valid in the presence of cardinality and weight constraints.

CHAPTER 4

GENETIC ALGORITHM APPROACHES

Heuristics are used for difficult problems that are not practical to be solved to optimality. Instead of optimal solutions, heuristics aim for satisfactory solutions that can be obtained more easily with short computation times. PO problems get more difficult as we consider additional and/or nonlinear criteria, more investment options and constraints that lead to integer or binary variables. Heuristics have been used to handle such PO problems. Genetic algorithms are a family of heuristic search techniques where a population of abstract representations evolves toward better solutions in successive generations. In genetic algorithms, we start with an initial population of feasible solutions. Each solution is represented by a set of chromosomes. From this population, parents that perform well with respect to the criterion/criteria of the problem are selected as parents to produce the offspring. To produce the offspring, we first apply crossover to the parents. Exchanging chromosomes between parents, we obtain children that contain properties from both parents. This is performed with the hope of obtaining better solutions than the parents. After the crossover, we also apply mutation to some of the children created. The purpose here is to change some properties of the offspring so that we can obtain solutions that can bring differentiation to the population. After the mutation, we select best out of the starting population and the offspring, and treat the selected solutions as the new population for the next generation. This process is repeated for a number of generations, evolving toward better solutions.

In this chapter, we first apply a widely-known genetic algorithm, Nondominated Sorting Genetic Algorithm II (NSGA-II) by Deb et al. (2002) to PO. The performance of NSGA-II is evaluated with two and three-criteria PO. In the second part of the chapter, we develop our genetic algorithm to handle a reference point-based, maximum cardinality and maximum weight-constrained PO problem. This algorithm is applied with expected return and variance criteria.

4.1 NSGA-II Approach

4.1.1 Review of NSGA-II

Basic ideas of NSGA-II are to use a non-dominated sorting method to assign fitness to solutions, and also to employ a crowding distance to attain diversified solutions. General steps of NSGA-II appear below (see Deb et al., 2002 for details):

1. Generate an initial population of size *L*.
2. Apply non-dominated sorting to the population and assign fitness values according to the domination fronts found.
3. Apply tournament selection to the population. To determine the winner, first consider fitness values; if two members have the same fitness, look at crowding distances.
4. Perform crossover (single-point crossover for binary-coded genetic algorithms and simulated binary crossover for real-coded ones) and mutation (bitwise mutation for binary-coded genetic algorithms and polynomial mutation for real-coded ones) and produce offspring of size *L*.
5. Combine the population and the offspring and apply non-dominated sorting again.
6. Accept domination fronts to the next generation until *L* slots are filled. In the last front that cannot be fully accepted (if any), choose based on crowding distance.
7. Continue with the next generation from Step 3.

The crowding distance of NSGA-II is designed to ensure diversity among generated solutions. While non-dominated sorting brings elitism, crowding distance tries to make sure that all regions of efficient frontier are represented. After domination fronts are found, crowding distances are used to compare solutions that are on the same front. The crowding distance of a solution (for two objectives) is half the perimeter of the rectangle formed by taking one solution on either side of that solution as the corners. The solution with the larger crowding distance is preferred.

4.1.2 Results of Experiments

We represent portfolios by strings of length equal to the number of available stocks. Each gene represents the proportion of a stock in the portfolio. With *n* assets available, a solution is represented as in Figure 14, where x_i is the proportion of stock *i*. Genes have real-valued variables. However, to ensure that constraint (2.3) is satisfied, x_i'*s* are normalized to sum up to one after initializing the population and after mutation.

x_1	x_2	x_3	…	x_n

Figure 14. Genetic representation of a portfolio

For our experiments, we use the NSGA-II code supplied by Kanpur Genetic Algorithm Laboratory (KanGAL)[3] with the default crossover and mutation operators.

[3] http://www.iitk.ac.in/kangal/codes.shtml, the original version

4.1.2.1 Expected Return–Variance Setting

We solve the model given by (2.1)–(2.3) in this two-criteria setting. We first experiment with 10 stocks from ISE to obtain preliminary results and gain insight (see Table 18 in Appendix A for the list of stocks). Their monthly return data between January 2007 and June 2008[4] are used to estimate expected return and variance values. We observe NSGA-II to perform well with 1000 generations, a population size of 35, 0.9 crossover probability and 0.1 mutation probability. Figure 15 illustrates solutions of NSGA-II against exact efficient solutions found by the augmented ε-constraint method. We see that NSGA-II performs well with respect to both convergence and diversity.

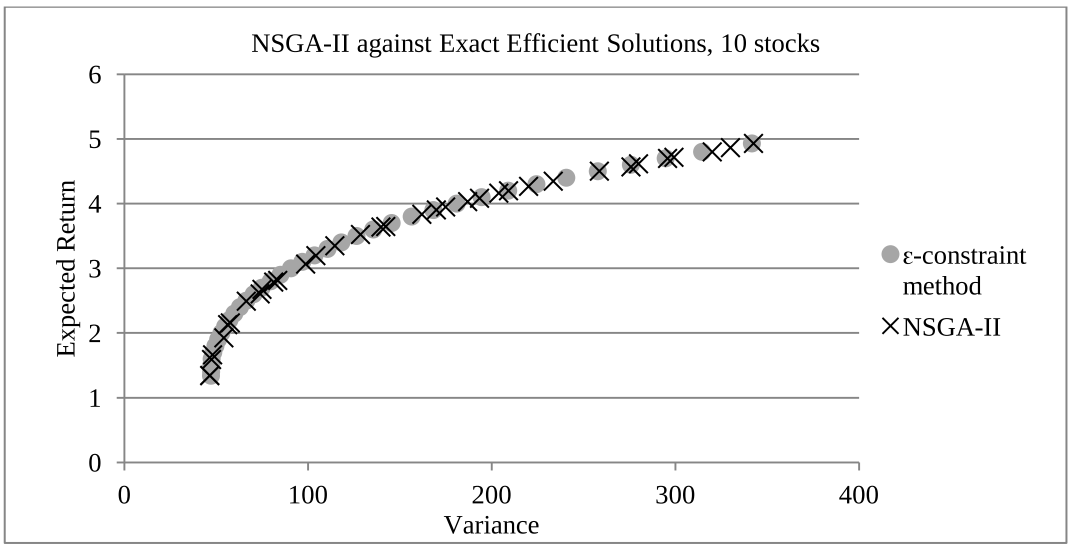


Figure 15. Expected return–variance solutions of NSGA-II and ε-constraint method with 10 stocks

To see the performance of NSGA-II in a larger problem, we next experiment with 100 ISE stocks, again using their monthly returns between January 2007 and June 2008 as input data[4]. Table 19 in Appendix A has the list of stocks. NSGA-II performs well with 1000 generations, a population size of 100, 0.9 crossover probability and 0.1 mutation probability. Assessing the performance of NSGA-II against 100 exact efficient solutions, we see that NSGA-II performs very well with respect to convergence. It also represents most parts of the efficient frontier, but misses the lower tail with the lowest expected return and variance values. As a remedy, we use end points of the exact efficient frontier as seeds in NSGA-II. The diversity performance of the heuristic has improved. Figure 16 illustrates the comparison of seeded NSGA-II to exact efficient solutions. Several performance metrics are used to evaluate the performance of

[4] http://borsaistanbul.com/en/data/data/equity-market-data/equity-based-data

evolutionary algorithms. Some evaluate the closeness to the efficient frontier and some evaluate the diversity among the solutions generated. Set coverage metric and generational distance (see Deb, 2001, p. 311-313) are examples of the former type; and spacing and spread (see Deb, 2001, p. 313-316) are examples of the latter. There are also metrics that evaluate both convergence and diversity. To provide a measure of the performance of NSGA-II, we use a metric of this kind: hypervolume. Hypervolume measures the dominated portion of the objective space by a set of solutions. Using the nadir point as the reference point, the hypervolume of the efficient frontier of the augmented ε -constraint method is 545.8409 whereas it is 543.3281 for NSGA-II. The ratio of the hypervolume of NSGA-II to exact solutions is 0.9954, which implies very good performance.

We also compare portfolio compositions of NSGA-II and exact solutions and observe that they are similar. The stocks contained in the portfolios and their proportions are alike.

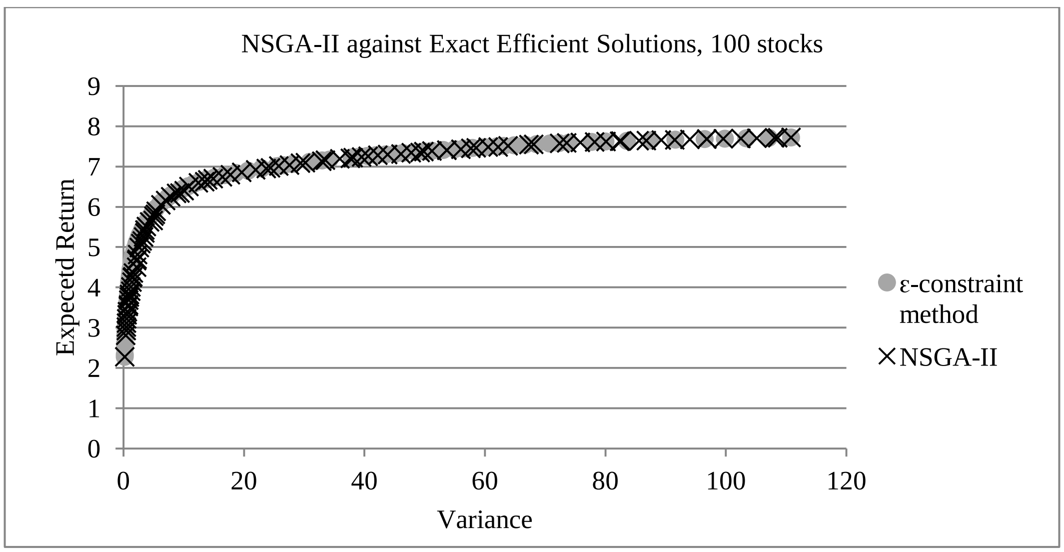


Figure 16. Expected return–variance solutions of NSGA-II and ε-constraint method with 100 stocks

4.1.2.2 Expected Return–Variance–MUA Setting

In this section, we again work with the 100 stocks introduced in Section 4.1.2.1 and their corresponding data. We want to experiment with an additional risk measure besides expected return and variance. We initially consider MAD, one of the most widely-used risk measures. We first make tests to see if variance and MAD are suitable measures to be used together. To see if they result in similar efficient portfolios, we find the corresponding MAD values for efficient expected return–variance points and the variance values for efficient expected return–

MAD points. Figures 45 and 46 in Appendix B show that the two risk measures behave similarly, so that the inclusion of MAD as the third criterion is not meaningful.

In search of another meaningful risk measure, we consider linear risk measures that are only interested in losses. We choose to utilize MUA that is defined by (2.25) in Section 2.3.3.3. It is a linear risk measure of expected returns below a threshold. For our experiments, we set 1.5% as the threshold. We compare expected return–variance and expected return–MUA efficient portfolios in expected return and variance axes in Figure 47 in Appendix B. We see that MUA can bring differentiation to our problem.

With three criteria of expected return, variance and MUA, we again run NSGA-II with 1000 generations, 0.9 crossover probability and 0.1 mutation probability; these parameters are observed to result in good solutions. The 100 solutions generated are plotted against the 100 solutions of expected return–variance NSGA-II in Figure 17. The corresponding MUA values of the expected return–variance NSGA-II solutions are calculated for this purpose. We can again see that the third criterion brings differentiation to solutions.

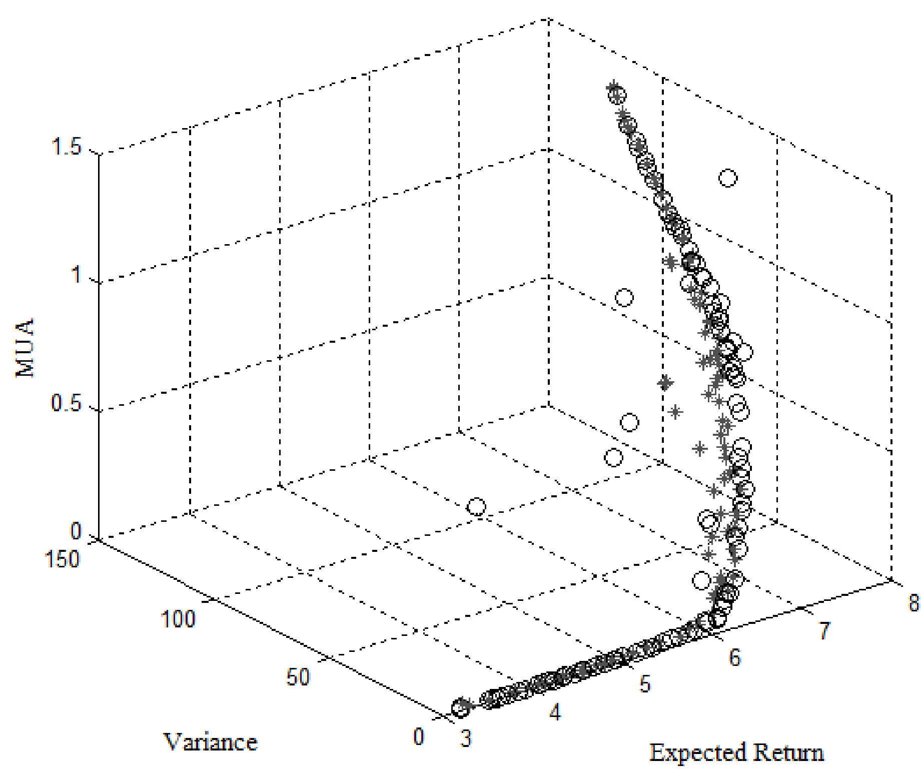


Figure 17. Comparing expected return–variance NSGA-II (o) and expected return–variance–MUA NSGA-II (*) solutions in three criteria

To further justify the use MUA as the third criterion, we check for portfolios in three-criteria NSGA-II that have comparable expected return and variance values to portfolios in expected return–variance NSGA-II, but considerably better MUA values. Table 3 includes such example solutions where we can improve MUA greatly with small sacrifices from expected return and/or variance. The differences in criteria's values are given in percentages with respect to the observed range of each criterion in expected return–variance NSGA-II solutions. We conclude that the addition of our third criterion results in considerably more valuable solutions for a DM who wants to take under-achievements into account.

Lastly, we want to assess the performance of 3-criteria NSGA-II against exact solutions. With the augmented ε-constraint method, we use the expected return and variance values of the three-criteria NSGA-II as constraints and optimize MUA. Working with a number of representative solutions, Figure 18 compares the results of NSGA-II against the ε-constraint method.

Table 3. Comparing solutions of two and three-criteria NSGA-II in expected return (ret), variance (var) and MUA

	Expected return–Variance NSGA-II			Expected return–Variance–MUA NSGA-II					
Portfolio pair	ret	var	MUA	ret	% difference	var	% difference	MUA	% difference
1	5.16	3.27	0.90	5.20	**0.97**	3.85	**0.53**	0.00	**-62.46**
2	5.88	6.03	1.36	5.93	**1.03**	7.31	**1.16**	0.00	**-93.57**
3	6.05	6.78	0.18	6.05	**-0.06**	8.57	**1.63**	0.00	**-12.26**
4	6.61	13.02	0.57	6.61	**0.04**	14.66	**1.49**	0.21	**-24.99**

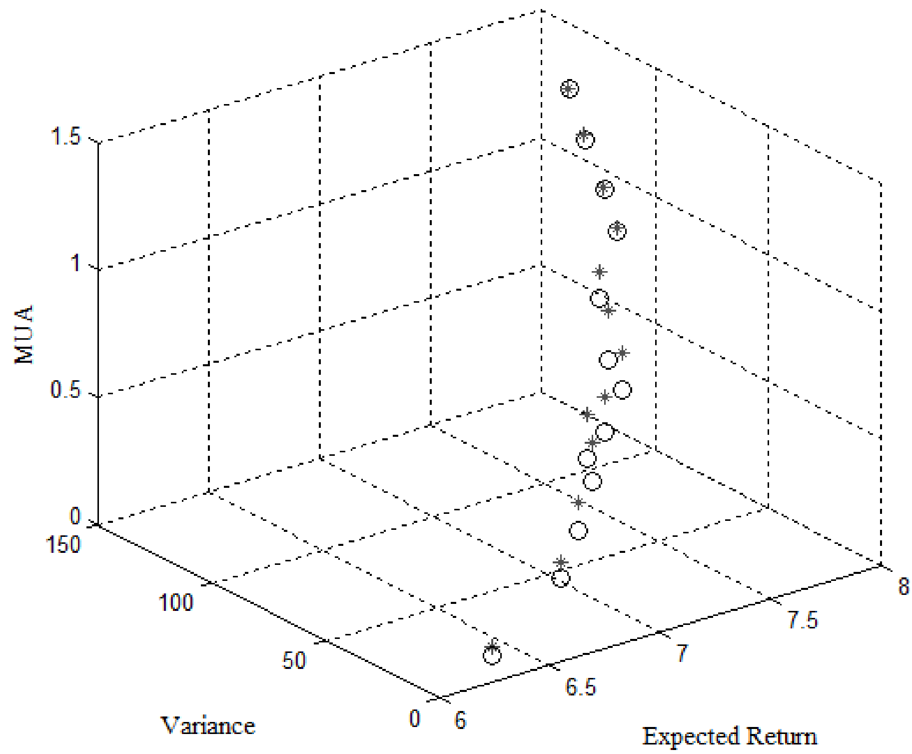


Figure 18. Comparing expected return–variance–MUA NSGA-II (*) and ε-constraint method (o) solutions

4.2 A Preference-Based Genetic Algorithm to Solve Portfolio Optimization with Cardinality and Weight Constraints

Looking at the literature of heuristics for PO with complicating constraints, we see that Chang et al. (2000) studied mean-variance PO with minimum proportion and cardinality constraints. They applied a genetic algorithm, a tabu search method and a simulated annealing approach. When dealing with cardinality constraints, they formulated their genetic algorithm to handle instances where the number of securities is fixed to a given number. They claimed that if they can handle this case, then they can run their algorithm for all values between the lower and upper bounds on the number of securities. For the fitness function, they incorporated the return criterion into the risk criterion with a weight. Children were generated by uniform crossover and two parents produced one child. Their mutation was limited to changing the weight of one security in the child. They employed a repair procedure as needed for the minimum weight constraint. They utilized a steady state population replacement strategy; in every generation, the child is replaced with the worst population member. Lin and Liu (2008) solved mean-variance PO with minimum transaction lots with genetic algorithms. PO with minimum transaction lots is a combinatorial problem which has a discontinuous feasible region. Minimum transaction lots require that security shares can only be bought in integers and usually in lots; one cannot purchase a fraction of a security and usually has to buy shares in multiples of some value. Lin et al. still represented solutions with chromosomes of real numbers for the sake of operational simplicity; later they found the maximum number of shares that can be bought with the given budget and prices. They used variance as the objective and handled expected return as a constraint, requiring that the minimum return should reach a specified value. Skolpadungket et al. (2007) also studied mean-variance PO with cardinality, floor and round-lot constraints with several well-known genetic algorithms, namely Vector Evaluated Genetic Algorithm (VEGA), Multiobjective Optimization Genetic Algorithm (MOGA), Strength Pareto Evolutionary Algorithm (SPEA) and NSGA-II. With these algorithms, they could treat the two criteria separately without reducing them to a single objective. Maximum cardinality, floor weights and round-lot restrictions were all handled as constraints. For all constraints, they used repair procedures.

Our method deals with a reference-point-based mean-variance PO problem with maximum cardinality and maximum weight constraints. Our multiobjective genetic algorithm handles both risk and return objectives simultaneously rather than trying to capture them in a single objective. In addition, as opposed to the study of Chang et al. (2000), it does not restrict the number of securities in the portfolio to a single value. Restricting the number of securities to a fixed value and running the algorithm for different such values is neither reasonable nor practical. The reasoning behind cardinality constraints is to limit transaction costs and to keep the portfolio manageable and observable. The most appropriate way to handle this is to allow different number of securities as long as a limit is not exceeded. Our algorithm handles this issue. On the other hand, it is known that risk reduction benefits are achieved when an investor holds a well-diversified portfolio. Since cardinality constraints impose upper bounds on the number of securities, they may prevent us from full diversification. As a result, we may want to limit our risk by other means; this is where our second constraint, the maximum weight constraint, steps in. We aim to limit our loss if any of our securities does poorly. Our algorithm can handle maximum weights that are greater than or equal to 0.5. In addition, we also take a reference point specified by the DM into account. Our algorithm directs the search towards a desirable direction for the DM, and tries to find solutions on the efficient frontier that are closest to the reference point.

The PO model handled by our algorithm is (2.1)–(2.3) extended by the following constraints:

$$x_i \leq b_i\, UB \qquad \forall\, i = 1, \dots, n \tag{4.1}$$

$$\sum_{i=1}^{n} b_i \leq K \tag{4.2}$$

$$b_i \in \{0,1\} \qquad \forall\, i = 1, \dots, n \tag{4.3}$$

where *UB* is the maximum weight allowed for assets and *K* is the maximum number of assets allowed in the portfolio.

We next give the details of our algorithm. We will see that it has specialized operators for the PO problem on hand. After the description of the algorithm, results of experiments follow.

4.2.1 The Genetic Algorithm

The genetic representation of portfolios is the same as in Section 4.1, which was illustrated in Figure 14. The general steps of our algorithm are given below, and they are explained in detail in subsections that follow.

Let $\bar{R}$ be the reference point in (expected return, variance) which takes normalized values [0,1]. Let x_i^j be the proportion of stock *i* in member solution *j*.

1. Ask the DM to determine *K*, *UB* and $\bar{R}$.
2. Generate an initial population of *L* feasible members.
   - Generate random solutions that satisfy constraint (4.2) and repair for constraint (4.1).
3. Find the objective function values and Tchebycheff distances of members from $\bar{R}$.
4. Apply tournament selection to determine the parents.
   - First check for domination. If the parents do not dominate each other, decide based on Tchebycheff distances.
5. Apply crossover to create offspring of size *L*.
   - Crossover operator is designed to satisfy (4.2). Repair for (4.1).
6. Apply mutation.
7. Calculate the objective function values and Tchebycheff distances of the offspring from *R*.
8. Make *L* pairwise comparisons between the initial population and the offspring. Determine *L* winners based on domination; and if there are ties, distance from *R*.
9. Carry *L* winners to the next generation as the initial population and repeat from Step 4 for a pre-set number of generations.

4.2.1.1 Initial Population Generation

For each population member, we generate a random integer *h* between 1 and *K*. We select *h* stocks from our security pool and assign random positive weights to the selected stocks. All other stocks assume zero weight. We next check for weight constraint (4.1). Since our algorithm is designed to handle *UB* values greater than or equal to 0.5, at most one stock will violate (4.1). Let *m* be the stock in solution *j* such that x_m^j > 0.5. Let $w = x_m^j$. Drop the weight of stock *m* to 0.5, i.e. x_m^j = 0.5. Let $T^j = \{n\text{: } 0 < x_n^j \leq 0.5\}$. Update the weights of stocks in T^j as follows: $x_n^j = x_n^j + \frac{(w-0.5)\, x_n^j}{1-w}$. This corresponds to allocating the excess weight of stock m to other stocks in the portfolio in proportion to their weights.

We should note at this stage of the algorithm that we see the cardinality constraint is satisfied with the random generation routine. Also note that we allow different member solutions to have different number of securities; thus we can represent different parts of the solution space with initial solutions and have enhanced exploration abilities.

4.2.1.2 Finding the Objective Function Values and Distances of Members from the Ideal Point

After finding the objective function values of solutions, we find their Tchebycheff distance from $\bar{R}$. The smaller this distance, the more preferable a solution is. Before we calculate the distances, we normalize expected return and variance values according to the following formula where v_i corresponds to the value to be normalized:

$$v_i = \left(\frac{v_i - v_{minimum}}{v_{maximum} - v_{minimum}}\right) \tag{4.4}$$

We do not calculate the exact ideal and nadir points in our algorithm, we use approximations. The same approximated ideal and nadir points are used for all parameters of the cardinality and weight constraints. The maximum and minimum expected returns are taken as the expected return of stocks with the maximum and minimum expected returns, respectively. For variance, the optimal variance is calculated with no cardinality and weight constraints, and it is used as the minimum variance value. And for the maximum variance value, the variance of the stock with the largest such value is used.

The normalized objective function values are used only for distance calculation purposes. Our algorithm continues to store and represent solutions with their original return and variance values.

4.2.1.3 Tournament Selection

From *L* population members, we form *L* pairs of competitors for parenthood. We randomly select two members from the population *L* times and make them compete based on two measures: domination and distance from $\bar{R}$. If a member solution dominates the other, it is directly assigned as a parent. Conversely, if there is no domination, the member with the smaller distance from $\bar{R}$ is assigned as a parent. The winner of the first pair becomes parent 1, the one from the second pair becomes parent 2 and so on, so that we select and label *L* parents.

4.2.1.4 Crossover

We apply crossover between parents 1 and 2, 3 and 4, … , *L*-1 and *L*, and obtain offspring of size *L*. With our crossover operator, we want to preserve enough characteristics of both parents in the offspring. We use two parents to obtain two children. To illustrate our crossover operator, let ns_i be the number of stocks with positive weights in parent *i*. Let us assume we are working with parents *j* and *j*+1. We set the number of stocks to be exchanged between the parents as: $\left\lfloor \frac{\min\{ns_j, ns_{j+1}\}}{2} \right\rfloor$, and select this number of stocks with positive weights from parents *j* and *j*+1. Then offspring *j* (*j*+1) is formed by copying the weights of selected stocks from parent *j*+1*(j)* and the weights of all remaining stocks from parent *j* (*j*+1). At this stage, a complication may occur. If any of the randomly selected stocks from parent *j* (*j*+1) has zero weight in parent *j+1(j)*, the number of stocks in offspring *j*+1*(j)* will exceed ns_{j+1} (ns_j). This may result in a violation of the cardinality constraint, which we do not want to encounter. To prevent this in a way that will promote reachability, we perform the following routine: After the crossover exchange, if the cardinality constraint is violated in offspring *j* (*j*+1), then the weights of stocks carried from parent *j* (*j*+1) are set to zero starting from the first stock until the cardinality constraint is satisfied. On the other hand, even if the cardinality constraint is not violated, the weight of each stock in offspring *j* (*j*+1) that is carried from parent *j* (*j*+1) is still set to zero with a probability. This enables us to achieve offspring with varied number of stocks.

After crossover is finalized, we normalize the weights to sum up to one. If we violate the maximum weight constraint, we repair with the procedure explained in Section 4.2.1.1.

4.2.1.5 Mutation

For each offspring, we perform mutation with a probability. In the offspring to be mutated, we randomly select two stocks. Our only selection criterion here is that at least one stock should have positive weight. The offspring is then mutated by exchanging the weights of the selected securities. There are two possible outcomes of this routine: weights of stocks with positive weights will be exchanged, or one stock will be removed from the portfolio by giving its weight to a previously left-out stock. These two outcomes enhance reachability of solutions and also do not violate the cardinality constraint.

4.2.1.6 Offspring Evaluation

After the offspring are generated, we calculate their objective function values and Tchebycheff distances from $\bar{R}$. These will be used in selecting the new population for the next generation.

4.2.1.7 New Population Selection

By this step, we have the original population and the newly-formed offspring, both of size *L*. The next generation can be determined by employing non-domination sorting and accepting members with the best fronts and minimum distances, but this can cause premature convergence. Instead, we choose to utilize a tournament selection procedure here as well. We make pairwise comparisons between population member *i* and offspring *i*, for *i* =1 to *L* with the procedure defined in Section 4.2.1.3. We end up with *L* new population members for the next generation. The algorithm is run for a specified number of generations.

4.2.2 Results of Experiments

Our algorithm is tested with two kinds of settings: First we run our algorithm with no cardinality and weight constraints. This can easily be handled by setting *K* to the number of available stocks and *UB* to one. We study if our algorithm can converge the efficient frontier in the region pointed by the reference point $\bar{R}$. We discuss this case by observing the true efficient frontier and the solutions of the algorithm on graphs. Secondly, we run our algorithm with cardinality and weight constraints, and assess its performance by calculating the deviations of generated solutions from exact efficient points. We make tests with instances of three different problem sizes. Using random expected returns and random valid covariance matrices, we experiment with 25 and 50 stocks. Then we also experiment with 100 actual stocks from ISE.

In our experiments, we make tests with different genetic algorithm parameters to see with which our algorithm performs well. We try different numbers of generations, population sizes, and crossover and mutation probabilities; and report our findings for the combination that our algorithm performs best with.

4.2.2.1 Tests with 25 and 50 stocks with Random Covariance Matrices

For both 25 and 50 stocks, we use random expected return vectors and random valid covariance matrices generated by the method of Hirschberger et al. (2007). The expected returns used for the 25 and 50 stock cases are given in tables 20 and 21 in Appendix A; and the covariance matrices for the two cases are given in tables 22 and 23 in Appendix A.

First with 25 stocks, we set *K* = 25 and *UB* = 1. We try different reference points in this setting. The first reference point, (expected return, variance) = (1,0) is expected to guide the search towards intermediate regions of the efficient frontier. We expect the second reference point (0.5,0) to result in a region where we have lower variance and expected return values as opposed to (1,0). Lastly, (1,0.5) should converge the high expected return, high variance region. With 25 stocks, we consider 100 and 200 generations, population sizes of 10, 20, 30 and 40, 0.9 crossover probability, and 0.1 and 0.5 mutation probabilities. In this setting, our algorithm performs well with 200 generations, a population size of 20, 0.9 crossover probability and 0.1 mutation probability. Figure 19 illustrates the results. We see that the solutions of our algorithm and the exact efficient frontier converge well in the regions expected.

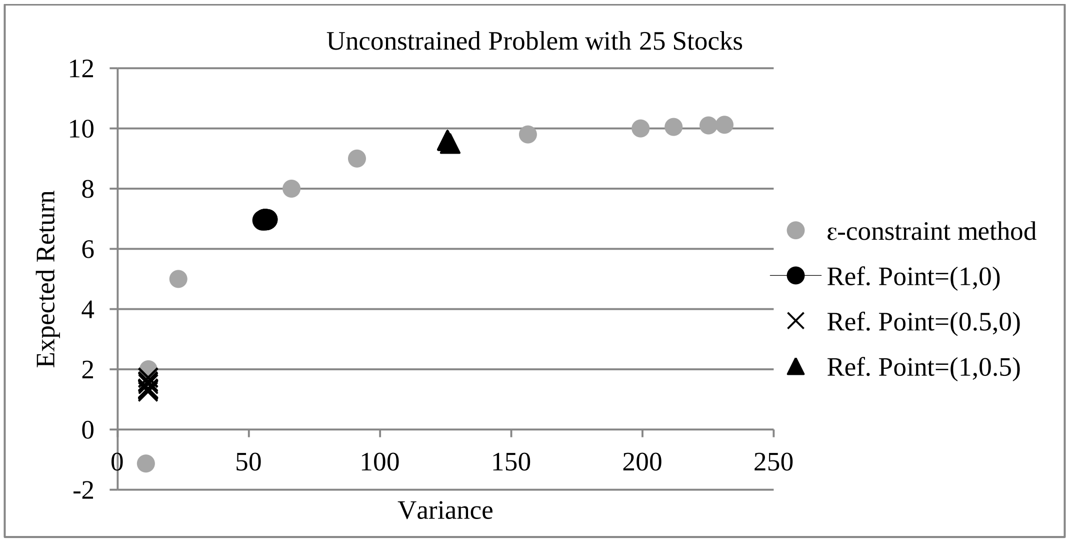


Figure 19. Solutions of the genetic algorithm proposed and ε-constraint method with 25stocks without constraints

Now we introduce cardinality and weight constraints with K = 6 and UB = 0.6. We take $\bar{R}$ = (1,0). In this setting, our algorithm is observed to perform well with 200 generations, a population size of 40, 0.9 crossover probability and 0.5 mutation probability. Looking at the resulting solutions, we see that they are formed in a quite concentrated region, i.e., they are very close to each other. Therefore we select two representative solutions and compare them to exact efficient solutions. For this purpose, we find the optimal variance values for the expected return values observed in the solutions, and compute the deviations of heuristic variances from optimal values. Table 4 shows the results; our algorithm has very small deviations.

Table 4. Comparing solutions of the genetic algorithm to efficient solutions with 25 stocks with constraints

Solution	Expected Return	Variance	Optimal Variance	% deviation
1	7.203	51.591	51.348	0.473
2	7.308	53.494	53.180	0.590

We continue with an instance of 50 stocks without cardinality and weight constraints, i.e., K = 50 and UB = 1. We use (1,1) as $\bar{R}$ and expect it to lead us to the region of high expected returns and variances. With 50 stocks, we consider 100 and 200 generations, population sizes of 20 and 40, 0.9 crossover probability, and 0.1 and 0.5 mutation probabilities. In this setting, our algorithm performs well with 200 generations, population size of 20, 0.9 crossover probability and 0.1 mutation probability. Figure 20 illustrates that our algorithm converges the efficient frontier in the expected region.

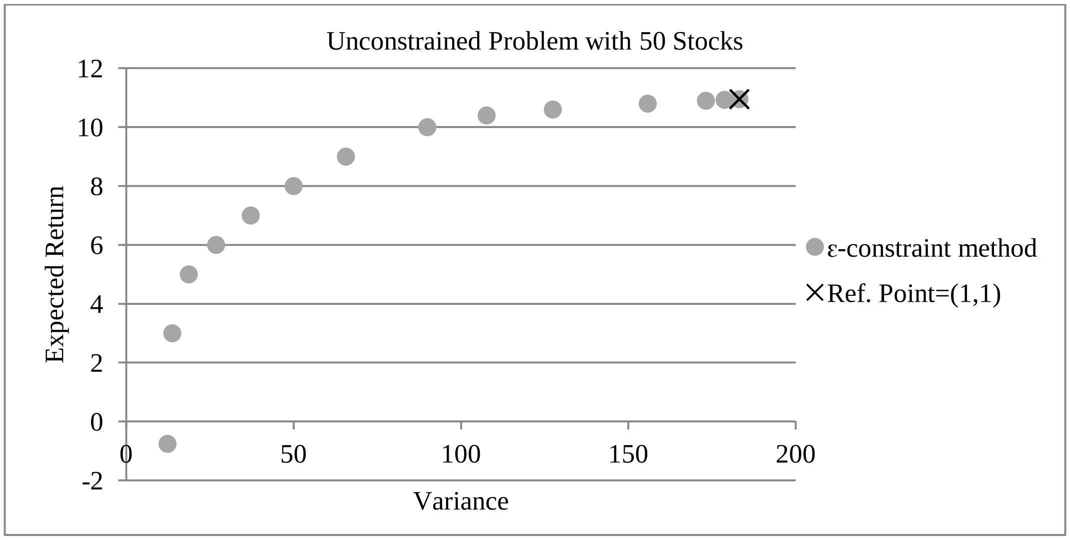


Figure 20. Solutions of the genetic algorithm proposed and ε-constraint method with 50stocks without constraints

We then set K = 10 and UB = 0.5, and $\bar{R}$ = (1,0). The best results are achieved with 200 generations, a population size of 20, 0.9 crossover probability and 0.1 mutation probability. We see that the solutions obtained are very close; in fact, they are not practically different from each other. Hence we take one of them as a representative solution and we find the optimal variance for its expected return. Table 5 shows that the deviation from optimal variance is again very small.

Table 5. Comparing a solution of the genetic algorithm to an efficient solution with 50 stocks with constraints

Solution	Expected Return	Variance	Optimal Variance	% deviation
1	10.44	23.272	23.167	0.453

4.2.2.2 Tests with 100 Stocks from ISE

As stated before, for our 100-stock case, we use actual stocks from ISE. The same set of stocks and their corresponding data used in Section 4.1.2.1 for NSGA-II tests are used here as well. With 100 stocks from ISE, we consider 200, 500 and 1000 generations, population sizes of 50 and 70, 0.9 crossover probability, and 0.1 and 0.5 mutation probabilities. We first set K = 100, UB = 1 and $\bar{R}$ = (1,0). Our algorithm performs well with 500 generations, a population size of 50, 0.9 crossover probability and 0.5 mutation probability. Figure 21 shows that our algorithm and the efficient frontier converge well in the region pointed by the reference point.

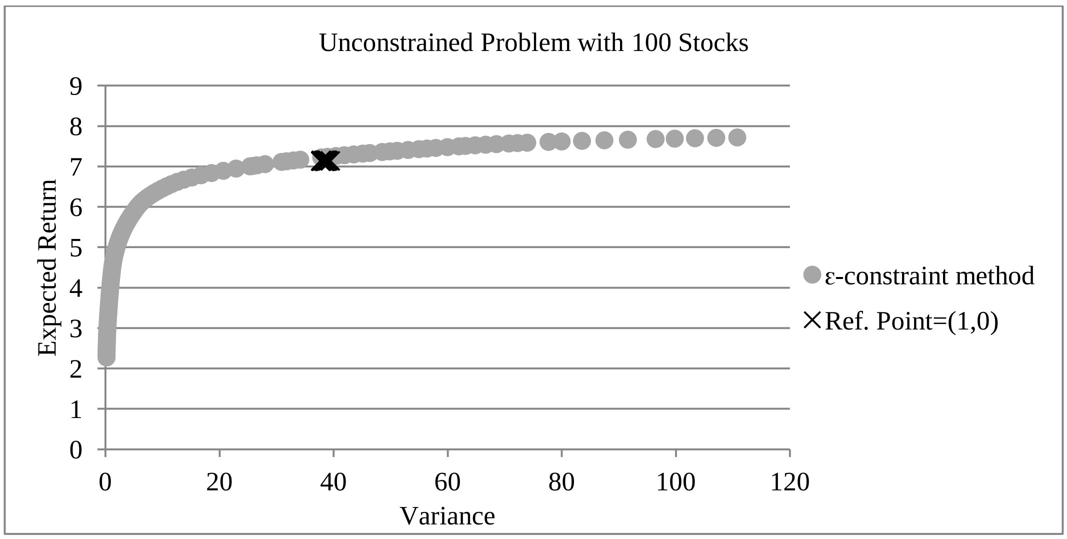


Figure 21. Solutions of the genetic algorithm proposed and ε-constraint method with 100 stocks without constraints

Lastly, we set K = 10, UB = 0.5 and $\bar{R}$ = (1,0). Our algorithm performs well with 500 generations, a population size of 70, 0.9 crossover probability and 0.1 mutation probability. The generated solutions are again concentrated in a limited region; hence we select two representative solutions to assess the heuristic performance. Table 6 shows the deviations of heuristic variances from optimal values. We see that the deviations are more pronounced in this case. However, these large deviations can in part be attributed to the wide range of variance in efficient solutions. Percentage deviations can be evaluated in the efficient range of variance; which will result in lower values.

Table 6. Comparing some solutions of the genetic algorithm to efficient solutions with 100 stocks with constraints

Solution	Expected Return	Variance	Optimal Variance	% deviation
1	7.175	36.430	34.553	5.432
2	7.144	35.372	32.634	8.390

To conclude, we have developed a genetic algorithm that can handle PO with upper bounds on the number and weights of stocks. The algorithm has specialized crossover and mutation operators for the cardinality constraints, and a repair procedure for the weight constraint. We have carried out experiments with 25, 50 and 100 stocks. Our algorithm was observed to perform well with respect to convergence with 25 and 50 stocks. With 100 stocks, the performance of the algorithm deteriorated to an extent; although the generated solutions were still close to the exact efficient solutions from a practical point of view.

We can further justify our algorithm with the help of computational studies on the run times of the heuristic and the exact method. This will be studied as future work. There are also some issues that need to be attended to improve our algorithm. Currently, we ask the DM for a single reference point and direct our search towards that point. As a result, we achieve a very concentrated region where the obtained solutions are sometimes not practically different. If the DM is not satisfied with the produced region, we will have to run the algorithm with different reference points. Instead, we can direct our search towards a region in some range of the reference point. For this purpose, we may work with a small set of reference points positioned around the point selected by the DM. In the tournament selection, if the solutions do not dominate each other, we can select the one that is closest to a point in the reference point set and this point can be determined probabilistically for each selection. We can also consider an interactive approach. Instead of collecting the DM preferences before the process, we can conduct our search with the help of the DM. This will increase the likelihood of the DM's satisfaction with the proposed solutions. As another issue, we can increase the robustness of our algorithm by updating its design to handle any *UB* value. We can also develop and test different crossover operators in search of a more effective one. Such issues will be attended as future work.

CHAPTER 5

THE STOCHASTIC PROGRAMMING APPROACH

Most of the studies in the literature on portfolio optimization have traditionally considered a single-period. With time, SP has come forward as an approach to handle multi-period PO where uncertainty of the future is explicitly accounted for. This uncertainty of the future is a result of the behavior of random parameters that affect the portfolio outcome. If these random parameters can be represented by discrete distributions, then SP can work with scenario trees. These scenario trees model the future movement of economical factors by assigning probabilities to different possible outcomes. A general scenario tree starts with a set of initial possible outcomes for the first period; and then continues with future periods with possible outcomes conditional on the realization of the previous periods. When all periods of the planning horizon are accounted for, we obtain a scenario path for each chain of realizations with given probabilities. The scenario tree has nodes from which possible branches grow. The DM is expected to make decisions at these nodes evaluating the information available on the future evolution of factors. The goal is to maximize final expected prosperity. We provide an example SP scenario tree in Section 5.2.2.

Abdelaziz et al. (2007) proposed a new deterministic formulation to multicriteria SP by combining compromise programming and chance constrained programming models for PO. The criteria they considered are rate of return, liquidity measured as exchange flow ratio and the risk coefficient. They applied their method with 45 stocks from the Tunisian stock exchange. Ibrahim et al. (2008) studied single-stage and two-stage SP models with the objective of minimizing maximum downside deviation. They used past returns of stocks as equiprobable scenarios. Yu et al. (2004) used SP in the bond market. They used a dynamic model with the objective of maximizing the difference between the expected wealth at the end of the investment horizon and the weighted sum of shortfall cost. Gülpınar et al. (2003) considered transaction costs with four asset classes, a number of liabilities and riskless assets. They minimized risk for given wealth levels and made tests with different numbers of scenarios. Pınar (2007) developed and tested multistage portfolio optimization models using a linear objective composed of expected wealth and downside deviation from a target. He used a simulated market model to randomly generate scenarios. Balibek and Koksalan (2010) developed a SP model for multi-objective public debt management problem. Yu et al. (2003) provided a survey on SP models in financial optimization. Their study starts with an introduction to SP and they discuss SP models for asset allocation problem, fixed-income security management and asset/liability management.

Making use of scenarios on market conditions, our multicriteria SP approach provides us with investment decisions for future periods. In this chapter, we present our work centered around this approach, with the exception of the interactive approach of weighted Tchebycheff programs that will be covered in Chapter 6. In Section 5.1, we discuss the conditions and assumptions of financial markets since we generate the scenarios accordingly. Then in Section 5.2, we structure our scenario generation technique and introduce our basic SP model. We use expected return, liquidity and CVaR as criteria in this model. We perform tests using stocks from ISE, present our major findings and share our insights. Sections 5.1 and 5.2 draw upon our paper Tuncer Şakar and Köksalan (2013b). In Section 5.3, we consider generating large numbers of scenarios to represent the market better. To keep the problem manageable, we use clustering on the generated scenarios. Section 5.4 encompasses our tests to see the effects of adding fixed-income securities and stocks that move in the opposite direction of the market to the available asset pool. We provide theorems, corollaries, remarks and examples on the properties of optimal SP solutions with different criteria and time periods in Section 5.5. Our purpose is to formalize the behavior of SP models for some contexts and justify the use of multi-period models over single-period ones. Lastly, in Section 5.6 we consider rolling horizon settings in SP. The revision of scenarios after certain economical factors are realized is evaluated with expected return and CVaR criteria.

5.1 Market Structure and Assumptions

The behavior of financial markets has been the subject of heated debates over the years. Academicians, finance companies and investors have tried to understand and predict the movement of markets. Since our scenario generation technique is related to our conclusions about the Turkish market, in this section we look into the hypotheses and assumptions about the financial markets. As a result, we develop the model we will use to represent the Turkish Stock Market (TSM).

5.1.1 Efficient Market and Random Walk Hypotheses

The question of whether financial asset prices can be predicted accurately is directly related to the hypotheses of market efficiency. There are three forms of market efficiency: weak, semi-strong and strong. Weak form market efficiency asserts that given the past price and volume information on financial assets, it is impossible to know future prices. Semi-strong form claims that with past information and all publicly available information (including all financial statements of companies disclosed to the public, news and expectations about the companies and the economy in general), it is still not possible to predict future prices. The last form of efficiency, strong form, argues that even with all possible public and private information, including insider information about companies, there is again no room for abnormal profits. There have been many studies testing the market efficiency hypotheses. Earliest studies were on testing weak-form market efficiency. Lo and MacKinley (1988) studied weekly returns of New York Stock Exchange and observed serial positive correlation over short periods. However, the correlation coefficients were rather small, suggesting significant abnormal profits were not possible. The study of Poterba and Summers (1988) for longer periods showed negative serial correlation; but it was claimed that the underlying reason may be varying risk premiums rather than market inefficiency. Concerning tests of semi-strong efficiency, there are findings that suggest some factors, such as a stock's price-earnings ratio and market capitalization, can predict abnormal returns; for example, see Basu (1983). However, Fama and French (1993) explained the short and long-term serial correlations between asset returns and the return-predictive nature of such factors as signs of risk premiums. The studies are more consistent for strong form efficiency. Several studies including Seyhun (1986) showed that insiders can trade profitably on their stocks, rejecting strong form market efficiency. To conclude, the most consistent finding is that strong form efficiency is not practically possible. But the evidence is mixed for the other two forms; there have been studies confirming both the existence and invalidity of efficient markets. Nevertheless, we consider the most prominent result to be that, even if there are some opportunities for beating the market, the time, energy and cost required to exploit these opportunities level off abnormal profits.

Random walk models are closely related to market efficiency and they are sometimes used as an approach to test it. See Campbell et al. (1997) for models and tests of random walk hypotheses. They directly follow from the assumptions of efficiency. If all related information about financial asset prices, including expectations, are already reflected in prices, then the change in the price of an asset tomorrow can only be random. If not, players in the market could have exploited their knowledge and earned extra returns.

Tests of random walk models for U.S. markets offer mixed results too. A full consistency cannot be expected since they use different time periods, financial markets and test methods. Bodie et al. (2009) discuss the efficient market hypotheses, their tests and random walk models in detail. They analyzed several findings considering possible underlying explanations.

The situation for Turkish markets is also debatable. It is common to accept that Turkish financial markets do not offer as much liquidity and competition as U.S. markets. As a result, Turkish markets are not expected to be on the same level of efficiency with U.S. markets. Even so, there have been studies in favor of market efficiency in Turkey as well. Smith and Ryoo (2003) carried out tests of random walk hypothesis for emerging European markets including Turkey. They studied the ISE-100 index and failed to reject the random walk hypothesis. Buguk and Brorsen (2003) tested weak-form market efficiency for composite, financial and industrial indices in Turkey and found that all three types of indexes obey random walk hypothesis. Odabasi et al. (2004) found that, for periods not longer than a year, monthly returns of ISE-100 follow a random walk. They also suggest that asset prices in ISE have evolved in time towards a more efficient state as the financial markets developed. Acknowledging studies in favor of efficiency of Turkish markets (weak-form), we think that the TSM can be assumed to follow a random walk model. To reinforce this assumption, we also study a TSM-representative index, ISE-100, to test for signs of patterns in its return.

Lo and MacKinlay (1988) tested the random walk hypothesis for several indexes and portfolios by comparing variance estimators of data sampled at different frequencies. Although they rejected the random walk hypothesis for weekly returns of different indexes and portfolios, they could not reject it for indexes with monthly return. Conrad and Kaul (1988) used autocorrelation to test time variation in expected returns. They showed that correlation coefficient gets smaller as portfolio size and the lag increase. They generally found significant correlations. We utilize the autocorrelation of ISE-100 returns to test if the index follows a random walk. Given measurements Y_t at time t, t=1, 2, …, n, the lag-k autocorrelation function is defined as:

$$r_k = \frac{\sum_{t=1}^{n-k}(Y_t - \bar{Y}).(Y_{t+k} - \bar{Y})}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2} \quad (5.1)$$

We want to test if autocorrelation of our index differs from zero and we are interested in lag-1 observations. We utilize two approaches for this purpose. First, we construct 95% confidence intervals for autocorrelation employing t test with test statistic $\hat{\hat{r}}/\sqrt{\frac{1-\hat{r}^2}{n-2}}$. Second, we test the significance of the lag coefficients of autoregression equation at 95% confidence level. For monthly returns between January 2003 and June 2010, we fail to reject zero autocorrelation with both approaches. These results support the efficiency of the index, but we also make tests with half-daily session returns to see if there are temporary correlations. Using data from January to March 2012, zero correlation cannot be rejected for session returns as well, confirming market efficiency further.

To keep our SP model manageable, we will derive individual stock returns from the ISE-100 index. The model we will use for this purpose is the Single Index Model and it is explained in the next section.

5.1.2 The Single Index Model and Our Random Walk Model

The Single Index Model relates the return of assets to a common macroeconomic factor. Using historical data, it uses regression analysis to regress the return of assets on the return of the macroeconomic factor. Since we are dealing with the stock market, the ISE-100 index will be a valid proxy for that factor. The regression equation for the Single Index Model is:

$$R_i(t) = \alpha_i + \beta_i\, R_M(t) + e_i(t) \quad (5.2)$$

Here $R_i(t)$ and $R_M(t)$ are the return of security i and the market-representative index at time t, respectively. β_i is called the beta of the security, it is the sensitivity of the security to the index. α_i is the security's return when the market return is zero; thus, it is an individual premium and its value is expected to be zero when security prices are in equilibrium. Lastly, e_i is the zero-mean, security-specific surprise in the return of the security.

When we take the expected value of the above equation, we obtain

$$E(r_i) = \alpha_i + \beta_i\, E(r_M) \quad (5.3)$$

where $E(r_i)$ and $E(r_M)$ are the expected return of security *i* and the index, respectively.

As a result, if we regress a stock's return on the return of ISE-100, we can estimate its α and β values. We use raw return values in our Single Index Model, that is, they are not adjusted to represent the excess return over the risk-free rate. If we used adjusted returns, we would expect α's to be zero in market equilibrium; stocks should not consistently offer more return than the risk-free return rate. However, in our approach, non-zero values of α are expected and we use them in our applications.

The Single Index Model enables us to approximate stock returns from a market index. We will refer the return of a stock estimated based on the market index as its "estimated return". On the other hand, the return obtained from our SP model is based on scenarios weighted by their likelihoods. Therefore, we will refer to this return as the "expected return."

The random walk model we will use to model the return of the ISE-100 index is as follows:

$$R_t = \mu + \sigma e_t \qquad e_t \sim N(0,1) \tag{5.4}$$

where R_t is the return of the index and e_t is a standard normal error term at time t, μ is the mean and σ is the standard deviation of the return of the index. When the mean return is non-zero, we have a random walk model with drift. We use a constant variation as it is supported by empirical evidence from ISE shown in Figure 22 where we present the time series plot of ISE-100 returns observed at half-daily sessions for work days from 2005 to 2010.

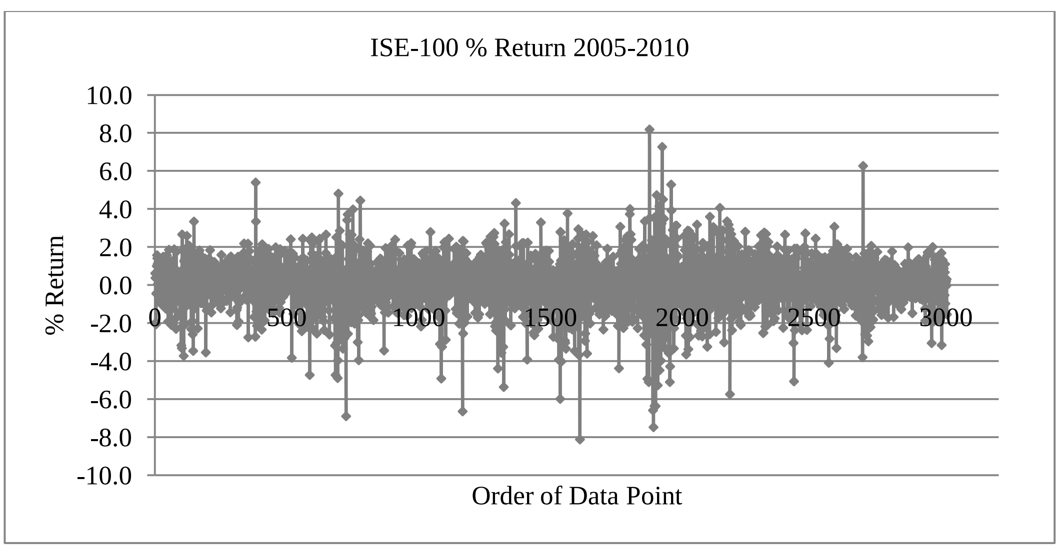


Figure 22. Session returns of ISE-100 for period 2005-2010

Normal distribution is frequently used in the literature for asset returns. From Figure 23, we observe that the return distribution of the ISE-100 index from 2005 to 2010 is in line with the applications in the literature and we are justified using (5.4) to generate returns for the ISE-100 index.

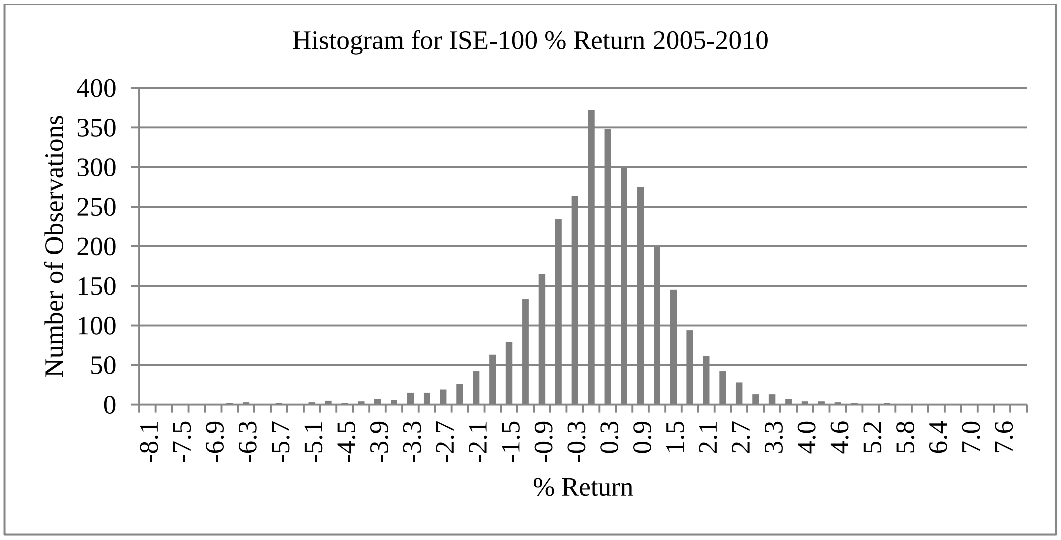


Figure 23. Histogram of % session returns of ISE-100 for period 2005-2010

5.2 Basic Stochastic Programming Approach to Portfolio Optimization

As discussed previously, SP is concerned with modeling optimization problems in the presence of uncertain parameters. It has been widely used to model multi-period asset management problems subject to uncertain economic factors. See Balıbek (2008) for an introduction to SP and an application to a multi-objective public debt management problem. In the following subsections, we first review scenario generation techniques for SP, especially for financial models. We provide the basics of our scenario generation technique. Then, we present our SP model and results of experiments.

5.2.1 Scenario Generation

Scenario generation is a fundamental part of SP and several methods to generate scenarios are proposed and implemented in the literature. Yu et al. (2003) discussed three methods to generate scenarios. The first, bootstrapping historical data, selects random historical occurrences of asset returns. The second method uses time series analysis of historical data to estimate volatilities and correlation matrices among assets. After the parameters are estimated, they can be used in Monte Carlo simulations. The third set of methods they discussed are vector autoregressive models that capture the progress of multiple time series while accounting for the interdependencies between them. Dupacova et al. (2000) discussed methods to generate scenario trees for multi-stage stochastic programs. They discussed random walk models, binomial and trinomial models and autoregressive models. Limiting their study to single-period PO, Guastaroba et al. (2009) considered historical data technique, bootstrapping technique, block bootstrapping technique, Monte Carlo simulation techniques and multivariate generalized ARCH process technique. They found that historical data technique gives solutions close to other techniques despite its simplicity. Hoyland and Wallace (2001) proposed a method that generates a limited number of discrete scenarios that possess DM-specified statistical properties. They implemented their method with single and three-period scenario trees.

In our studies, we need to generate scenarios on the stocks that we use in our model. Generating scenarios on individual assets would make our scenario tree very complicated and increase the computational difficulty substantially. Therefore, as stated earlier, we derive estimated stock returns from the ISE-100 index by the Single Index Model (5.3). We generate scenarios on our index using the random walk model (5.4).

5.2.2 The Model

Our basic SP model has three criteria: expected return, CVaR and liquidity. After the mathematical model, we explain the constraints and give the model size in terms of decision variables and constraints. But first, we provide a hypothetical scenario tree in Figure 24 that will aid the reader in following the model. The tree has three stages and two branches for each decision node, resulting in $2^3 = 8$ scenarios. Node 3 is the final node of scenario 1, node 2 is the immediate predecessor of node 3, and nodes 0, 1, and 2 form the predecessor set of node 3.

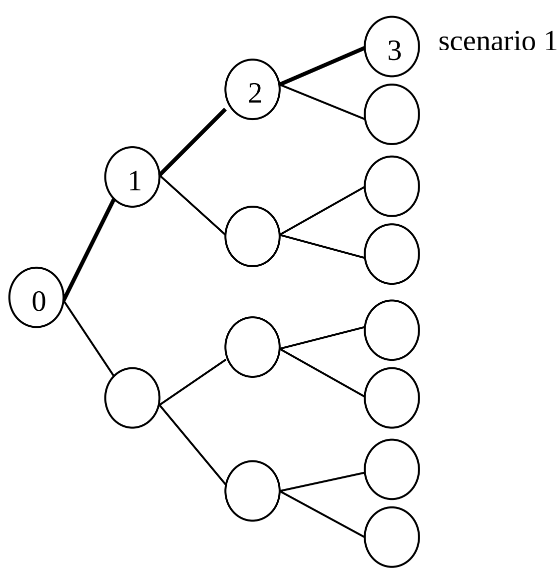


Figure 24. An illustrative 3-period, 15-node, 14-branch scenario tree

Parameters:
I: set of all stocks
S: set of all scenarios
N: set of all nodes in the scenario tree
n_s: the final node of scenario $s \in S$
n': immediate predecessor node of $n \in N$
$PR(n)$: set of all predecessor nodes of $n \in N$
$m(n', n)$: market return corresponding to branch between nodes n' and n
$p(n', n)$: probability corresponding to branch between nodes n' and n
p_n: probability of the partial scenario up to node n, where $p_n = p_{n'} . p(n', n) \quad \forall\, n \in N$, n = 0 is the root node of the scenario tree, and $p_0 = 1$.
liq_i: liquidity value of stock $i \in I$
α_i: stock-specific return premium of stock $i \in I$
β_i: sensitivity of stock $i \in I$ to the market
λ = probability level for CVaR
D = number of periods in the scenario tree

Decision Variables:
VaR: Value at Risk value
x_{ni}: allocation of stock $i \in I$ at node $n \in N \backslash \{n_s : s \in S\}$

Structural Variables:
Ret_s: value obtained if scenario s is realized, $s \in S$
$liquidity_s$: resulting liquidity if scenario s is realized, $s \in S$
aux_s: auxiliary variable corresponding to scenario $s \in S$

$$\text{Max } z_1 = \sum_{s \in S} p_{n_s} . Ret_s \tag{5.5}$$

$$\text{Min } z_2 = VaR + \left[\frac{1}{1-\lambda} \sum_{s \in S} (p_{n_s} . aux_s) \right] \tag{5.6}$$

$$\text{Max } z_3 = \sum_{s \in S} p_{n_s} . liquidity_s \tag{5.7}$$

$$\sum_{i \epsilon I} x_{ni} = \sum_{i \epsilon I} x_{n'i} (1 + \alpha_i + \beta_i . m(n', n)) \quad \forall\, n \in N \backslash (\{0\} \cup \{n_s : s \in S\}) \tag{5.8}$$

$$Ret_s = \sum_{i \epsilon I} x_{n'_s i} (1 + \alpha_i + \beta_i . m(n'_s, n_s)) \quad \forall\, s \in S \tag{5.9}$$

$$aux_s \geq 1 - Ret_s - VaR \quad \forall\, s \in S \tag{5.10}$$

$$liquidity_s = \frac{\sum_{k \in PR(n_s)} (\sum_{i \in I} x_{ki} . liq_i)}{D} \quad \forall\, s \in S \tag{5.11}$$

$$\sum_{i \in I} x_{0i} = 1 \quad (5.12)$$

$$x_{ni} \geq 0 \quad \forall\, n \in N \backslash \{n_s : s \in S\}, i \in I \quad (5.13)$$

$$aux_s \geq 0 \quad \forall\, s \in S \quad (5.14)$$

For the expected return criterion, we maximize the compounded return over all periods. To obtain a linear model, we use balance equations for decision nodes. At the initial node at time zero, we start with an initial wealth value of 1. At the other nodes, we allocate the ending value of a period among the stocks selected for the next period. With (5.12), we choose the initial percentages of stocks. (5.8) ensures that we allocate the ending value of each node except the final stage nodes to stocks chosen at its successor node. (5.9) gives the values of final nodes, thus the ending values of scenarios. With (5.5), we maximize the expected value of the scenario tree; and since we have started with an initial wealth of 1, z_1–1 gives the compounded expected return over all periods. (5.6), (5.10) and (5.14) are adopted from the linear CVaR model discussed in section 2.3.3.2. (5.10) and (5.14) makes use of auxiliary variables to keep the excess loss of each scenario beyond the VaR level, and (5.6) gives the CVaR value. In (5.11) we calculate the weighted sums of the liquidity measures of the stocks for nodes of each scenario, and then use their average as the liquidity measure corresponding to that scenario. (5.7) gives the liquidity criterion in which we take the weighted sum of the liquidity measures of scenarios by their corresponding probabilities. In sign restrictions of (5.13), we exclude the final nodes from the set of stock allocation decision variables since these nodes are the ending points of scenarios and no decisions are made at those points.

If CVaR is negative in the optimal solution of our model, it means that given the probability level for CVaR, the expected value of worst results is still a gain over our initial value.

The size of the model depends on the number of stocks, scenarios and branches, and the number of nodes in the scenario tree. The number of nodes increases as the number of branches in the scenario tree increases. The decision variables consist of stock allocation decisions for all combinations of stocks and nodes except for the final nodes of scenarios; and a variable to represent *VaR*. As a result, the number of decision variables is $|I|.|N \backslash \{n_s : s \in S\}| + 1$. The total number of constraints (5.8) and (5.9) is equal to $|N|$-1. For each scenario, there is one constraint of type (5.10) and one constraint of type (5.11). Additionally, there is the initial stock allocation constraint for the root node, so the number of constraints equals $|N| + 2|S|$. As an example, the scenario tree in Figure 24 has 8 scenarios and a total of 15 nodes. For this tree, we would have $7|I|$+1 decision variables and 31 constraints.

5.2.3 Results of Experiments

First, we generate scenarios on the return of the ISE-100 index. We obtain the published historical monthly data on the closing price of the index from January 2003 to June 2010[5]. We

[5] http://borsaistanbul.com/en/data/data/equity-market-data/equity-based-data

then calculate the corresponding monthly return values. Using these, we estimate the mean and standard deviation of the index return as 2.29 and 9.80 percentage points, respectively. Then we choose 100 random stocks from ISE, and carry out regression analysis to estimate α_i and β_i values for each stock i. The most recent data available at the time of the application was December 2009 for the monthly percentage returns of the stocks. Therefore, we use the data from January 2003 to December 2009[6]. Out of 100 stocks, five of them fail to have significant β values at a 0.05 significance level. We exclude these and use the remaining 95 stocks in our applications. Table 24 in Appendix A has the list of stocks. Three of these stocks also have significant α values. For the liquidity measure, we use the daily average number of shares traded[6] and the total number of outstanding shares[7] corresponding to December 2010.

Before we solve our problem with three criteria, we first study pairs of criteria to observe how they interact and conflict. To find a discretized efficient frontier for both the bicriteria cases and the three-criteria case, we use the augmented ε-constraint method. For all our applications, we use the same 3-month scenario tree where each node (except the final nodes of scenarios) has 7 equiprobable branches, resulting in $7^3 = 343$ scenarios. For each branch of the scenario tree, we generate the return for the market index using the random walk model equation (5.4). The number of decision variables of our model can be calculated as $|I|.|N\backslash\{n_s: s \in S\}| + 1 = (95.57) + 1 = 5416$, where 57 corresponds to the number of nodes excluding the final nodes of scenarios. The number of constraints of our model turns out to be $|N| + 2|S| = 400+686 = 1086$. We also tried $5^3 = 125$ and $6^3 = 216$ scenario settings and results showed similar structures. Nevertheless, we present the results corresponding to the larger number of scenarios to have a more representative scenario tree.

5.2.3.1 Expected Return and CVaR

We prefer high expected return and low CVaR values; and we expect to see conflicting behavior between the two criteria. Figure 25 shows the efficient frontiers with different probability levels for CVaR. We can observe the tradeoffs between the two criteria in the figure. We use expected return in the objective function and CVaR in the constraint for the augmented ε-constraint method. We generate 10 efficient solutions for each probability level that are evenly spaced in the CVaR range.

Figure 25 exhibits some interesting properties. We observe that the CVaR values are realized at lower (better) values for low probability levels. This is expected since as we decrease the probability, we include better return values to our previous points, and their expectation (CVaR) improves. Another observation is that the minimum expected return realized on the efficient frontier is different for different probability levels. Maximum expected return is naturally the same for all probability levels since the upper bound of expected return is the same regardless of CVaR.

[6] http://borsaistanbul.com/en/data/data/equity-market-data/equity-based-data

[7] http://www.mkk.com.tr/wps/portal/MKKEN/InvestorServices/eDATACapitalMarketsDataBank

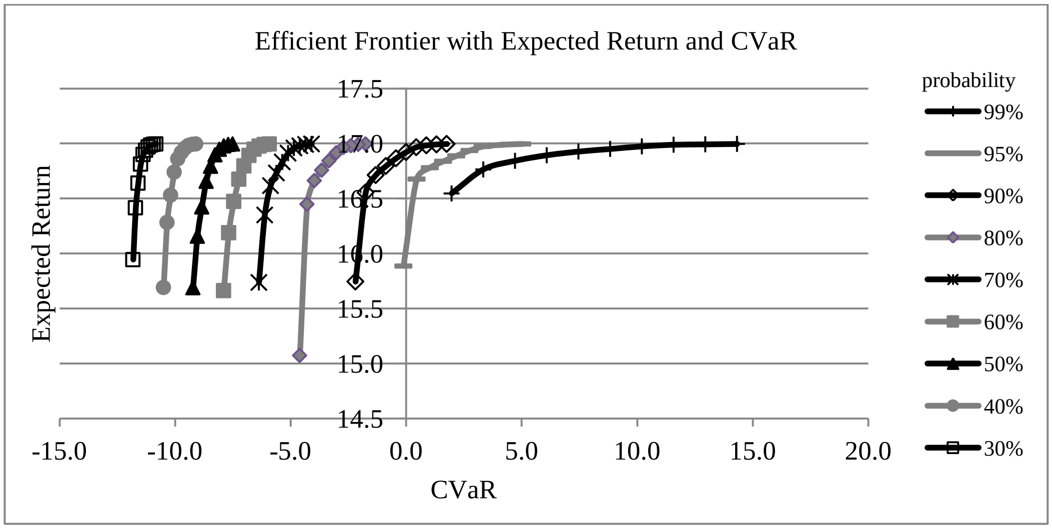


Figure 25. SP efficient frontier with expected return (%) and CVaR (%)

Figure 25 also reveals that the range for expected return is smaller for low and high probabilities, and larger for intermediate probabilities. This can be attributed to the realization that, for 100% probability level, CVaR will no longer be of concern to us, and we will be left with only the expected return objective. We will obtain a single solution that gives the maximum expected return. For probability levels close to 100%, expected return levels will be close to this maximum. At the other extreme, we have the case with 0% probability level where expected return and CVaR are identical. We will again find a single efficient point of maximum expected return, and as we increase our probability level, we will observe increasing tradeoffs between expected return and CVaR. So, for both endpoints of the range of CVaR probability levels, we achieve a single maximum expected return value, and at some point along the range, we realize a probability level which gives an efficient solution with the minimum expected return value. This probability is around 80% in our case.

Lastly, we see that for low probability levels, the sacrifice we have to make in CVaR to achieve higher expected return levels are not substantial. As the probability level increases, the range of CVaR also increases. This can be again attributed to the similarity of expected return and CVaR objectives for low probability levels. Considering 10% probability level as an example, CVaR considers the worst 90% returns over all scenarios for this probability, and one can make the analogy that expected return considers worst 100%. Thus, as we decrease the probability level, we observe positive correlation between the two criteria. As a result, optimizing for expected return also results in good values for CVaR. As we increase the probability level, conflicting nature of expected return and CVaR increases and the range for CVaR widens.

5.2.3.2 Expected Return and Liquidity

We prefer portfolios with high liquidity values. Investors would like to invest in stocks they can easily sell at their fair prices. Therefore, as previously stated, it is generally believed that illiquid stocks should offer extra return (liquidity premium) for their inconvenience in trading. Therefore, we expect to see a tradeoff between expected return and liquidity in our SP model. Figure 26 confirms our expectation where we generate efficient solutions by optimizing expected return using liquidity values as constraints in the ε-constraint model. We use 10 equally-spaced values within the efficient range of liquidity to generate 10 efficient solutions.

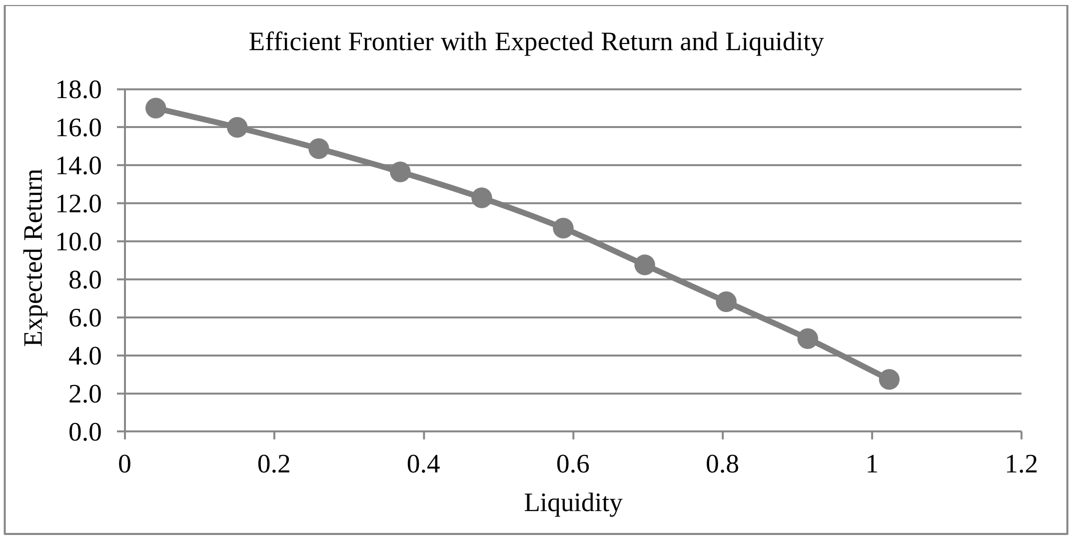


Figure 26. SP efficient frontier with expected return (%) and liquidity

5.2.3.3 CVaR and Liquidity

We have no intuitive expectation about the tradeoffs between CVaR and liquidity. Nevertheless, Figure 27 shows the conflicting nature of liquidity and CVaR with 90% probability level in SP settings. We find efficient solutions by optimizing expected return for 10 equally-spaced liquidity values. One explanation for the tradeoff could be that, risk is caused by changes in return values and such changes are affected by the trade volume of stocks.

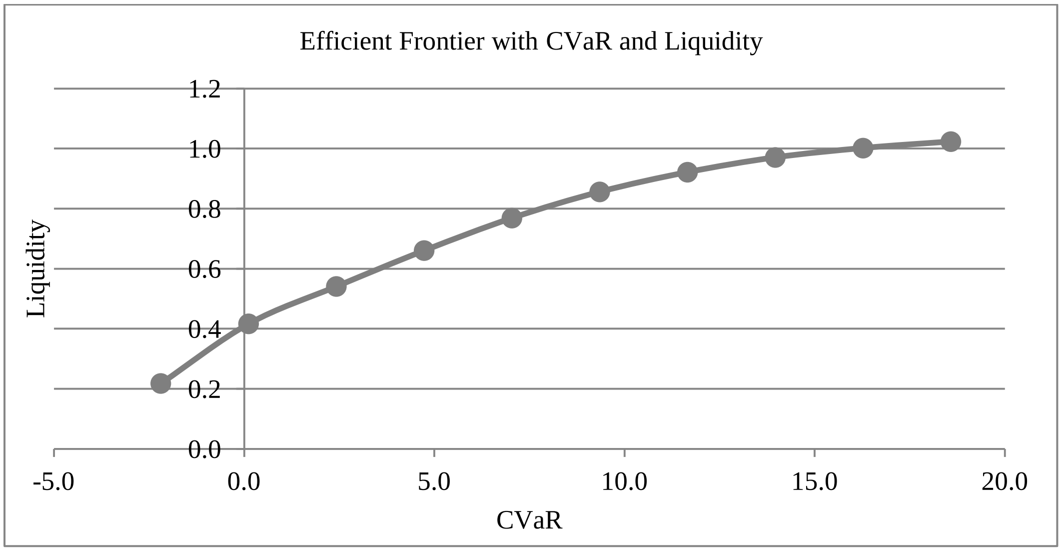


Figure 27. SP efficient frontier with CVaR (%) and liquidity

5.2.3.4 Expected Return, Liquidity and CVaR

In this three-criteria case, to apply the augmented ε-constraint method, we use the algorithm of Ehrgott and Tenfelde-Podehl (2001) to find the ideal and nadir points. Their algorithm gives the exact nadir point for discrete problems. Since we are utilizing a discretized efficient frontier, the nadir point we obtain is an approximation. We then optimize expected return for different levels of liquidity and CVaR. We apply a two-step procedure for this purpose. First we find 10 CVaR values that are equally spaced between its nadir and ideal values. Then for each of these 10 values, we find the range of feasible liquidity values and obtain 10 equally-spaced liquidity values. This gives us 100 pairs of CVaR and liquidity values to use in the augmented ε-constraint method. Next, we change the places of CVaR and liquidity and find another 100 pairs, obtaining 200 pairs in total to find the corresponding optimal expected return value.

Figure 28 illustrates the efficient solutions obtained, again with 90% probability level for CVaR. The ranges for expected return, liquidity, and CVAR are 2.74 to 16.99%, 0.04 to 1.02 and -2.19 to 18.86%, respectively. A careful analysis of Figure 28 reveals the tradeoffs between the three criteria at different regions of the solutions as well as the possible locations of efficient solutions. We observe that low levels of expected return are realized at the region of high liquidity values. High levels of CVaR are also realized for high liquidity values. Therefore, liquidity can be considered as the factor that forces expected return and CVaR to their worst values in this three-criteria case. This observation is also supported by the range of expected return and CVaR in the efficient frontier corresponding to the 90%-probability case of Figure 25. When we have expected return and CVaR as our criteria, their efficient ranges are 15.75 to 16.99% and -2.19 to 1.76%, respectively. We see that both criteria assume narrow ranges that contain good values. When the DM searches the solution space of the three-criteria case for solutions of interest to her/him, it would be helpful to take these types of characteristics into account. If she/he is interested in solutions that are good with respect to liquidity, she/he should be willing to sacrifice from expected return and CVaR. Another observation we can make from Figure 28 is that, although both expected return and CVaR exhibit substantial tradeoffs with liquidity, the lowest liquidity values appear when expected returns are at their highest values. As a result, if the DM is interested in the highest values of expected return, she/he should take into account that this region contains solutions that are rather good in CVaR but poor in liquidity.

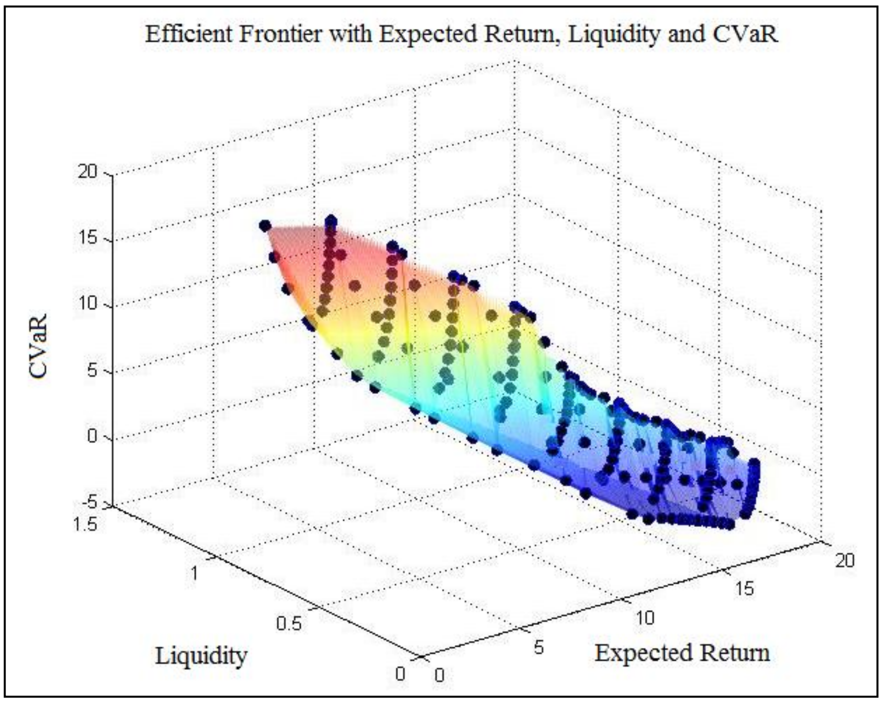


Figure 28. SP efficient frontier with expected return (%), liquidity and CVaR (%)

5.3 Clustering of Scenarios

As we have clarified by now, in our SP approach, we use scenario trees to represent the stochasticity of market movements. Using a random walk model, we create scenarios on the monthly return of our market- representative index and derive the returns of individual stocks from these scenarios.

Now, we want to increase the approximation capabilities of our scenario trees to represent the market behavior better. For this purpose, we consider generating a large number of scenarios;

but this has a drawback: computational complexity. Thus, we employ a two-stage procedure: first, generate a large number of scenarios, and then, classify them into clusters based on their similarity. This procedure is used in the SP approach to public debt management in Balıbek (2008).

Clustering is basically assigning items to groups so that they are more similar to the other items in the same group than the ones in other groups. Similarity between items is measured by a certain distance metric. There are several clustering methods, and the one we use is one of the simplest and most widely used methods, the *K*-means algorithm (MacQueen, 1967). The *K*-means algorithm clusters data points in a predetermined number (*K*) of disjoint classes. Each cluster has a centroid, which may or may not be one of the points in the clusters. The algorithm aims at minimizing the sum of distances of all points to the centroid. The steps of the *K*-means algorithm appear below:

1. Find a set of initial cluster centroids.
2. Assign each data point to the cluster to whose centroid it has the smallest distance.
3. After assigning all data points, recalculate the *K* centroids as the mean of points assigned to each cluster.
4. Repeat steps 2 and 3 until the number of point reassigned or the decrease in the sum of distances falls below some point.

To cluster our scenarios, we use the *K*-means function of the Statistical Toolbox of MATLAB. To find the initial cluster centroids, we employ the default setting of using *K* random elements from our data points.

We now provide our routine to obtain clustered scenario trees. Let t be the time counter, y be the number of periods in the scenario tree, c be the desired number of scenario paths stemming from each decision node, and M be the number of scenario paths to be clustered down to c.

1. Set $t = 0$, decide on the values of y, M and c.
2. For each node of time t, randomly generate M scenarios on ISE-100 return with the random walk model (5.4) by using random sampling from the error term.
3. Cluster each set of the M scenarios to c classes to form the scenario paths.
4. Compute the probability of each scenario path as the number of elements in the related cluster divided by M.
5. Set $t = t+1$ and repeat steps 2 to 4 until $t = y$-1.
6. Connect the scenario paths of time 0 to t to form final scenario paths covering all periods and compute their corresponding probabilities.

We generate clustered scenario trees that have equal number of branches coming out of each decision node since we make applications with such trees. Nevertheless, the above routine can easily be modified so that each node can have a different number of clustered paths.

To test the performance of our clustering approach in terms of representation capability, we use the measure proposed by Kaut and Wallace (2003) to test the stability of the generator. They assert that the optimal objective function values obtained by different scenario trees generated by a certain method should be distributed with a small variance. That is, the results should be stable across different replications.

We compare the stability of the results of clustered scenario trees against scenario trees that are formed without clustering using the same stock and index return data used in Section 5.2.3. We set y = 3, and c = 5, M = 100 for all n. We make 50 replications with both clustered scenario trees and trees that are formed without clustering. In each replication, we optimize for the expected return and CVaR criterion separately. We do not optimize for liquidity since it is independent of return realizations. We expect clustering to result in smaller variance values for both of our criteria. Table 7 shows the estimated mean and standard deviation of the expected return and CVaR criteria for scenario trees that are formed with and without clustering. Clustering the scenarios results in about a 76% decrease in the standard deviation of expected return, and a 68% decrease in the standard deviation of CVaR. Hence we achieve substantial improvement in robustness of scenarios by clustering them. See Balibek and Köksalan (2012) for a more thorough comparison of clustered and unclustered scenario trees. Using the public debt management problem, which can be considered a portfolio management problem as well, they made tests with different c and M values with y = 3. They also found that clustered scenario trees enhance stability.

Table 7. Stability results for unclustered and clustered scenarios

Scenario Trees Formed without Clustering			Scenario Trees Formed with Clustering		
Estimated	Expected Return (%)	CVaR (%)	Estimated	Expected Return (%)	CVaR (%)
Mean	20.7688	-3.6306	Mean	17.6820	-3.0376
St. Dev.	5.6893	3.0508	St. Dev.	1.3382	0.9755

Our SP applications for the TSM in the upcoming parts of the thesis utilize clustered scenario trees with y = 3, and c = 5, M = 100.

5.4 Effects of Negative Betas of Stocks and Fixed-income Securities

We use stocks to form our available asset pool for our SP approaches to PO. In Section 5.2.3, we made tests with expected return, liquidity and CVaR to see how they interact and conflict. We used stocks from ISE for this purpose, and all of the stocks used had positive betas (sensitivity coefficients to the market).

Now we want to study if negative betas for some stocks and the addition of a fixed-income security will bring differences to the behavior and conflict of criteria. For instance, will negative betas enlarge the range of CVaR? Will the fixed-income asset increase the maximum possible return of a portfolio? Fixed-income securities and stocks with negative betas can be advantageous when the market is going down.

For our tests in this section, we select 100 stocks from ISE and use their monthly percentage returns, and also the ISE-100 index return from January 2008 to December 2009[8]. The list of stocks is given in Table 25 in Appendix A. For the liquidity measure, we use the daily average number of shares traded[8] and the total number of outstanding shares[9] corresponding to June 2011. We use a 3-month scenario tree with 5 branches for each decision node, resulting in 125 scenarios. As discussed in Section 5.1.2, we estimate stock returns by the Single Index Model (5.3) by multiplying estimated market returns by market-sensitivity coefficients of stocks (betas) and adding the stock-specific premiums (alphas). We use regression analysis to estimate alpha and beta values. In our applications in Section 5.2.3, we only used stocks with statistically significant betas, and only employed significant alphas of these stocks. However, we think that a more reasonable procedure would be to use all beta and alpha values without significance tests. When we estimate the return of a stock from the market index, we directly use the estimated alpha and beta and do not include any error term. The insignificant alphas and betas are left out because their confidence regions include the value 0; but the variation of significant alphas and betas in their regions are unaccounted for. We believe this creates inconsistency, and continue our applications in this thesis by employing all alpha and beta values of the stocks. We do not revise our tests in Section 5.2.3 since they correspond to published material (Tuncer Şakar and Köksalan, 2013b).

5.4.1 Effects of Negative Betas of Stocks

We study the effects of negative betas using pairwise criteria of expected return, liquidity and CVaR. None of the 100 stocks considered has a negative beta as a result of the regression analysis. Throughout our studies covered in this thesis, we have not come across a stock with a negative beta value. Moreover, we have not encountered any studies in the literature that reported the existence of such stocks. Nevertheless, we want to see the potential effects of negative betas in a hypothetical setting. We initially find efficient expected return–CVaR, expected return–liquidity and CVaR–liquidity frontiers with the augmented ε-constraint method. Afterwards, we select some stocks and artificially negate their beta values. To select a representative set of stocks, we use a special case of a filtering method developed for vectors of values. With the Method of First Point Outside the Neighborhoods (see Steuer, 1986, p. 314-318), a set of vectors are filtered so that a number of vectors are retained that are different from each other. This method will be discussed in Chapter 6 as well; for now it suffices to say that it enables us to choose a number of betas that will be representative of the 100 ones generated by regression. We choose to select 10 betas to negate.

[8] http://borsaistanbul.com/en/data/data/equity-market-data/equity-based-data

[9] http://www.mkk.com.tr/wps/portal/MKKEN/InvestorServices/eDATACapitalMarketsDataBank

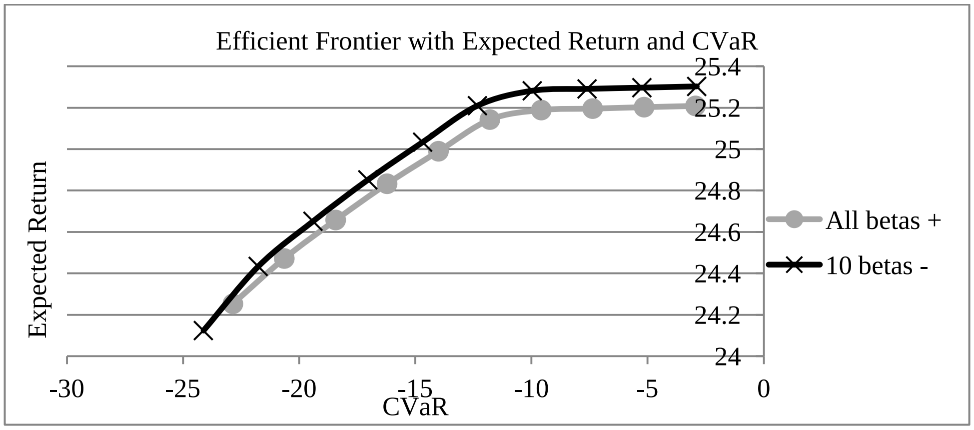


Figure 29. SP efficient frontier with expected return (%) and CVaR (%) with and without negative betas

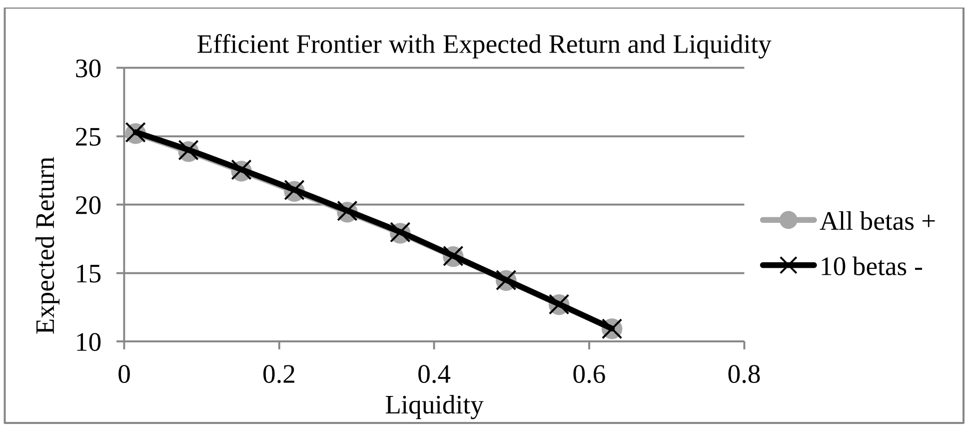


Figure 30. SP efficient frontier with expected return (%) and liquidity with and without negative betas

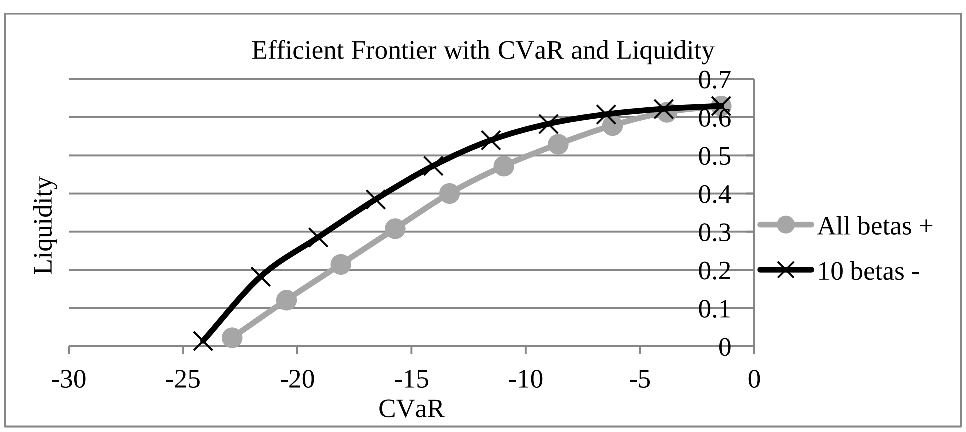


Figure 31. SP efficient frontier with CVaR (%) and liquidity with and without negative betas

With negative betas, we find the efficient frontiers with pairs of criteria and compare these to the original ones generated with real data. Figures 29, 30 and 31 illustrate the results. We see in Figure 29 that the expected return–CVaR efficient frontier with negative betas dominates that of the original data even if not drastically. We can deduce that negative betas became useful for scenarios with negative market returns. Both the maximum expected return and the minimum CVaR values improved. On the other hand, Figure 30 shows that there is no considerable difference for the pair of expected return and liquidity. The maximum liquidity level is of course independent of the beta values. The expected return values for given liquidity levels improved to an extent, though not considerably. Figure 31 shows that the CVaR–liquidity pair demonstrates the effect of negative betas more noticeably.

In our further tests that are not covered here for the sake of briefness, we see that the effects of negative betas would be more pronounced if we only used stocks with statistically significant betas and took their alphas only when they are significant too. In that case, with our data, none of the alphas turn out to be significant. As a result, the return is solely dependent on betas; and any change in them has more drastic effects.

5.4.2 Effects of a Fixed-income Asset

Again using the same pairs of criteria, we study the effects of adding a fixed-income security to the investment options. For our applications, we choose to employ a monthly bank deposit account for this purpose. We use 5% yearly interest rate effective at the date of the application. For the liquidity measure of the deposit account, we assume a value of one. This value is reasonable since our liquidity measure is like a turnover ratio and deposit accounts can be liquidated as soon as they mature.

We here want to clarify our use of data from different time periods for index and stock returns, stock liquidities and bank deposit account return. To calculate the parameters of the random walk and single index models, we use a period of historical returns to estimate the future index and stock returns. To obtain values that are more likely to be representative of the future, we use the most recent data available at the time of the application. For liquidity as well, we use data from the most recent month available to us. For the deposit account, we can obtain the most up-to-date interest rate effective at the time of the application, and we utilize this. We assume that this rate will be valid for the three months of our planning horizon. In brief, for all measures, we use the data that we assume to be representative for the planning horizon of our model.

We graph the efficient frontiers with the fixed-income asset against the case of only stocks in figures 32 and 33. We do not provide a graph of the expected return–CVaR case since the fixed-income asset has no effect on the efficient solutions. This is due to the fact that for all scenarios, there were stocks that were better alternatives than the deposit account. In Figure 32 we see that range of liquidity enlarged to include better values as expected. As stated previously, the deposit account has no effect on the maximum expected return value. However, the range of expected return enlarged to include lower values that occur as a result of including the deposit account in the portfolio. On the other hand, in the CVaR–liquidity case as illustrated in Figure 33, the worst CVaR value rose only slightly. The efficient frontier with the deposit account dominates the one with only stocks; we can obtain better liquidity values for given levels of CVaR.

Similar to the case of using negative betas, the effects of a fixed-income asset would be more drastic if we only used significant alphas and betas of stocks. Since none of the stocks have significant alphas, the stock returns would have no constant, and would be solely dependent on the market return. As a result, the addition of a fixed-income asset would affect expected return and CVaR more profoundly.

In our SP approach, the efficient solutions obtained depend on the scenario tree used model the future progress of the TSM. Different scenario trees result in different efficient frontiers. To account for this randomness, a statistical analysis may be employed. It is possible to work with multiple scenario trees to generate confidence ellipsoids around efficient solutions. Currently, to obtain each efficient solution, we optimize for one of the criteria while imposing bounds on the values of the others with the augmented ε-constraint method. We can repeat this process with multiple scenario trees for each solution, and use the theory on constructing confidence regions for multivariate means. These confidence regions can give the DM information about the probable variation to be expected in the criteria values of efficient solutions. This information can also be used to compare efficient frontiers with and without stocks with negative betas, and with and without fixed-income securities, as given in figures 29–33. The DM may study the overlaps of the confidence regions of solutions generated with two approaches, and see if the difference is statistically significant. The theory on constructing confidence regions for multivariate means is covered in Chapter 6 where we use it in our interactive approach to SP-based PO.

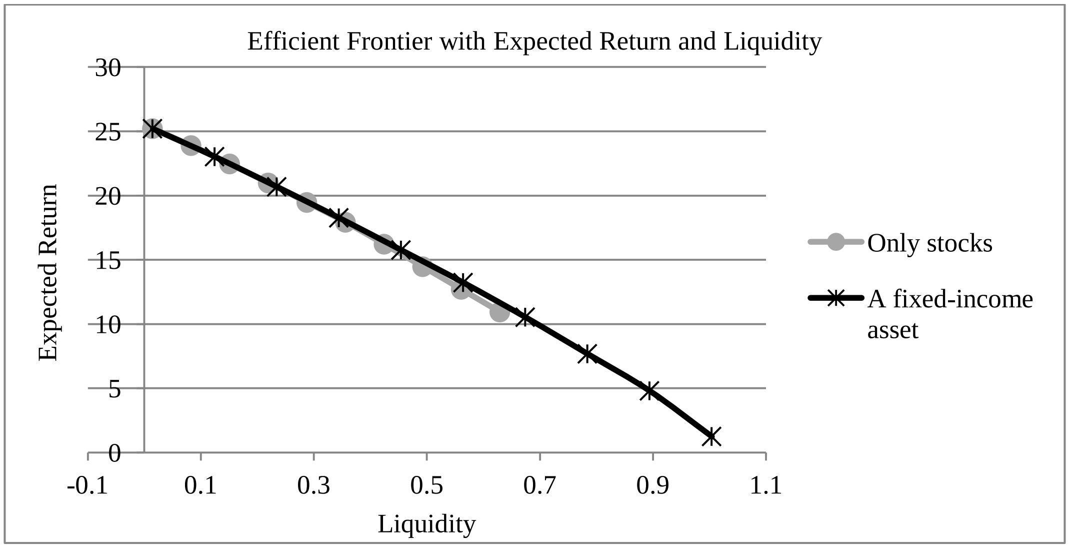


Figure 32. SP efficient frontier with expected return (%) and liquidity with and without a fixed-income asset

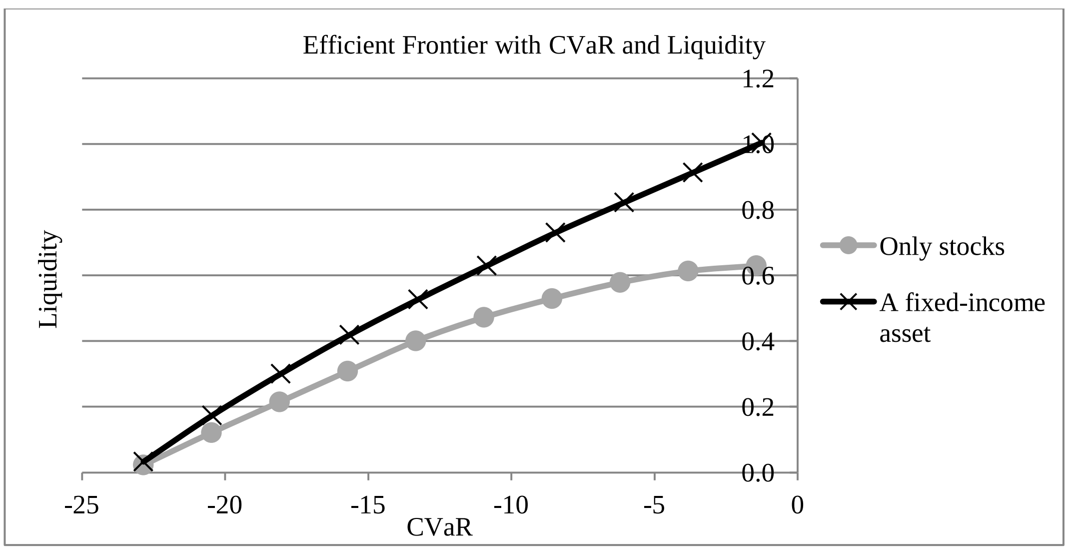


Figure 33. SP efficient frontier with CVaR (%) and liquidity with and without a fixed-income asset

5.5 Properties of Optimal Stochastic Programming Solutions with Different Criteria and Time Periods

In our previous work covered in Section 5.2, we made SP applications with three-period models using expected return, CVaR and liquidity criteria. We solved two and three-criteria models and provided insights into the behavior of criteria in various solutions.

In this section, we make a more formal analysis of the properties of optimal SP solutions. We look at single and multi-period models and provide comparisons for some contexts. Two factors make us question whether single-period models can be considered as substitutes for multi-period models: First, our PO problem is solved with stocks that can be traded in any decision node independently of the previous decisions. Thus, decisions may not be necessarily conditional on decisions of previous periods. Second, since we generate scenarios on market return utilizing random walk models, paths stemming from nodes of previous realizations are not dependent on those realizations. As a result, we want to test if we can solve individual single-period models and combine their solutions to obtain optimal multi-period decisions. This would bring us reduced computational complexity. However, we will see that multi-period models are likely to perform better when the DM has a planning horizon of multiple periods. We study different problems using expected return and CVaR in different settings. We exclude the liquidity criterion from our studies since it is independent of scenarios of market return. We discuss our findings in the context of the TSM where applicable.

We use the following notation:
I: the set of stocks
S: the set of scenarios
N: the set of all decision nodes
N_s: the set of decision nodes on scenario *s*
p_s: probability of scenario *s*
r_s: value of scenario *s*
a_s: auxiliary variable of scenario *s* to keep excess losses
m_s: market return of scenario *s* in single-period models
m_s^n :market return of scenario *s* on the path after node *n* in multi-period models
α_i: stock-specific return premium of stock *i*
β_i: market-sensitivity coefficient of stock *i*
x_i: proportion of stock *i* in single-period models
x_i^n: proportion of stock *i* in the portfolio of node *n* in multi-period models
VaR_λ: the variable that holds the value of VaR that is used to calculate CVaR
ρ : a sufficiently small positive constant

5.5.1 Maximizing Return in Single-period SP

Consider the single-period SP model of maximizing expected value of the portfolio:

$$max\ z = \sum_{s \in S} p_s . r_s \tag{5.15}$$

s. t.

$$r_s = \sum_{i \in I} x_i \,.\,(1 + \alpha_i + \beta_i . m_s) \quad \forall\, s \in S \tag{5.16}$$

$$\sum_{i \in I} x_i = 1 \tag{5.17}$$

$$x_i \geq 0 \quad \forall\, i \in I \tag{5.18}$$

(5.15)–(5.18) is referred to as *1-SP*$_{Ret}$.

Note that here we define r_s, the value of scenario *s*, as the ending value, which is equal to 1+return. As a result, the objective is to maximize the expected value of the portfolio, and this is equivalent to the objective of maximizing expected return. Owing to this equivalence, for convenience, we denote our model as *1-SP*$_{Ret}$ and generally use the expression return instead of value. We also denote the return criterion as *Ret*.

Theorem 1:
An optimal solution to *1-SP*$_{Ret}$ is $x_{i'} = 1$, $x_i = 0$ for $i \in I \setminus i'$ where $i' = \arg\max_{i \in I}(\alpha_i + \beta_i . \sum_{s \in S} p_s . m_s)$.

Proof:

$$max\ z = \sum_{s \in S} p_s . r_s = \sum_{s \in S} \left[p_s . \left(\sum_{i \in I} x_i (1 + \alpha_i + \beta_i . m_s) \right) \right] \tag{5.19}$$

$$= \sum_{i \in I} \left[x_i . \left(\sum_{s \in S} p_s . (1 + \alpha_i + \beta_i . m_s) \right) \right] \tag{5.20}$$

$$= \sum_{i \in I} \left[x_i . \left(\sum_{s \in S} p_s + \sum_{s \in S} p_s . \alpha_i + \beta_i . \sum_{s \in S} p_s . m_s \right) \right] \tag{5.21}$$

$$= 1 + \sum_{i \in I} \left[x_i . \left(\alpha_i + \beta_i . \sum_{s \in S} p_s . m_s \right) \right] \tag{5.22}$$

Since $\sum_{i \in I} x_i = 1$,

$$x_{i'} = 1 \text{ where } i' = arg\ \max_{i \in I} \left(\alpha_i + \beta_i . \sum_{s \in S} p_s . m_s \right) \quad \blacksquare \tag{5.23}$$

Corollary 1:

- If $\alpha_i = 0$ for all $i \in I$ and $\sum_{s\in S} p_s . m_s \geq 0$, an optimal solution to *1-SP_{Ret}* is $x_{i'} = 1$, $x_i = 0$ for $i \in I \setminus i'$ where $i' = \arg\max_{i\in I}(\beta_i)$.
- If $\alpha_i = 0$ for all $i \in I$ and $\sum_{s\in S} p_s . m_s \leq 0$, an optimal solution to *1-SP_{Ret}* is $x_{i'} = 1$, $x_i = 0$ for $i \in I \setminus i'$ where $i' = \arg\min_{i\in I}(\beta_i)$.

Proof: These are special cases and directly follow from Theorem 1. ■

Theorem 1 and Corollary 1 are applicable to any single-period portfolio problem of maximizing return, and thus can be used to directly see the optimal solution of $1\text{-}SP_{Ret}$ for the TSM.

5.5.2 Minimizing CVaR in Single-period SP

Consider the single-period SP model of minimizing CVaR:

$$min\ z = VaR_\lambda + \frac{1}{(1-\lambda)} \sum_{s\in S} p_s . a_s \tag{5.24}$$

s. t.

$$a_s \geq 1 - \sum_{i\in I} x_i . (1 + \alpha_i + \beta_i . m_s) - VaR_\lambda \quad \forall\, s \in S \tag{5.25}$$

$$\sum_{i\in I} x_i = 1 \tag{5.26}$$

$$a_s \geq 0 \quad \forall\, s \in S \tag{5.27}$$

$$x_i \geq 0 \quad \forall\, i \in I \tag{5.28}$$

(5.24)–(5.28) is referred to as $1\text{-}SP_{CVaR}$.

Let *f*(*m*) be the market return under different scenarios. Let $\psi(m, T)$ be the probability of *f*(*m*) not falling below a return level *T*. Then the *m-VaR* at λ probability level can be defined as $m{-}VaR_\lambda = max\{T \in R : \psi(m, T) \geq \lambda\}$.

Theorem 2:

If $\beta_i \geq 0$ for all $i \in I$, then an optimal solution to $1\text{-}SP_{CVaR}$ is $x_{i'} = 1$, $x_i = 0$ for $i \in I \setminus i'$ where $i' = arg\,max_{i\in I}\left(\alpha_i + \beta_i . \sum_{s\in S_{m-VaR_\lambda}} p_s . m_s\right)$ and S_{m-VaR_λ} is the set of scenarios where $m_s \leq m{-}VaR_\lambda$.

Proof:

Let q_{is} be the return of stock *i* under scenario *s*, $q_{is} = \alpha_i + \beta_i . m_s$.

If $\beta_i \geq 0$ for all $i \in I$, $q_{is} > q_{is'}$ will be true if and only if $m_s > m_{s'}$ is true. As a result, if we sort q_{is} values for any i for $s \in S$ in increasing order, it will be in increasing order of m_s. The worst (1- λ)% return values for any $i \in I$ will be realized under the worst (1- λ)% m_s values.

For any portfolio of multiple stocks, if we denote return under scenario s by Q_s, Q_s is given by:

$$Q_s = \sum_{i \in I} x_i \,.\, (\alpha_i + \beta_i . m_s) = \sum_{i \in I} x_i . \alpha_i + \sum_{i \in I} x_i . \beta_i . m_s \qquad (5.29)$$

Since $\sum_{i \in I} x_i . \beta_i \geq 0$ if $\beta_i \geq 0$ for all $i \in I$, sorted Q_s values in increasing order for any $\bar{x}$ will be in increasing order of m_s, and the worst (1- λ)% return values for any $\bar{x}$ will be realized under the worst (1- λ)% m_s values.

So, an optimal solution to *1-SP_{CVaR}* is:
$x_{i'} = 1$ where $i' = arg\,max_{i \in I} \left(\alpha_i + \beta_i . \sum_{s \in S_{m-VaR_\lambda}} p_s . m_s\right)$. ■

Corollary 2:
If $\beta_i \geq 0$ for all $i \in I$, and

- If α_i = 0 for all $i \in I$ and $\sum_{\mathrm{s \in S_{m-VaR_\lambda}}} p_s . m_s \geq 0$, an optimal solution to *1-SP_{CVaR}* is $\mathrm{x_{i'}}$ = 1, x_i = 0 for $i \in I \setminus i'$ where $i' = \arg\max_{i \in I}(\beta_i)$
- If α_i = 0 for all $i \in I$ and $\sum_{\mathrm{s \in S_{m-VaR_\lambda}}} p_s . m_s \leq 0$, an optimal solution to *1-SP_{CVaR}* is $x_{i'}$ = 1, x_i = 0 for $i \in I \setminus i'$ where $i' = \arg\min_{i \in I}(\beta_i)$

where S_{m-VaR_λ} is the set of scenarios where $m_s \leq m{-}VaR_\lambda$.

Proof: These are special cases and directly follow from Theorem 2. ■

Remark 1:
If $\beta_i \geq 0$ for some $i \in I$ and $\beta_{i'} < 0$ for some $i' \in I$, the worst (1- λ)% return values for i and i' will be realized under different m_s values. As a result, the optimal solution to *1-SP_{CVaR}* may involve a combination of multiple stocks.

Example 1:
Let us assume there are two stocks and four equiprobable scenarios.
$\bar{m}$ = (-0.2, -0.15, 0.1, 0.2), $\bar{\alpha}$ = (0.010, 0.015), $\bar{\beta}$ = (-0.4, 0.9).

When $\bar{x} = (1, 0)$, the value of the portfolio under different scenarios is given by $\bar{r}$ = (1.090, 1.070, 0.970, 0.930). With λ = 0.75, CVaR of this portfolio is realized as 1-0.930 = 0.070. When $\bar{x}$ = (0, 1), $\bar{r}$ = (0.835, 0.880, 1.105, 1.195). With λ = 0.75, CVaR of this portfolio is realized as 1-0.835 = 0.165.

Even if we randomly choose $\bar{x}$ = (0.5, 0.5), the value of the portfolio under different scenarios is realized as $\bar{r}$ = (0.9625, 0.9750, 1.0375, 1.0625); and with λ = 0.75, CVaR is equal to 1-

0.9625 = 0.0375, a better value than both of the cases above. It can be shown that the optimal CVaR with λ = 0.75 is -0.0115 with $\bar{x} = (0.692, 0.308)$.

We had not foreseen that the optimal solution to *1-SP*$_{CVaR}$ would consist of a single stock when $\beta_i \geq 0$ for all $i \in I$. When $\beta_i < 0$ for some i, this rule no longer applies. We did not have an intuition about this behavior before we studied the problem in detail.

In our applications, $\beta_i \geq 0$ for all ISE stocks considered. Therefore, Theorem 2 and Corollary 2 are applicable to ISE in general. This allows us to directly see the optimal solution to single-period models when the objective is to minimize CVaR in the TSM.

5.5.3 Maximizing Return in Single vs. Multi-period SP

Consider the multi-period SP model of maximizing expected return:

$$max\ z = \sum_{s \in S} p_s . r_s \tag{5.30}$$

s. t.

$$r_s = \prod_{n \in N_s} \left[\sum_{i \in I} x_i^n . (1 + \alpha_i + \beta_i . m_s^n) \right] \quad \forall\, s \in S \tag{5.31}$$

$$\sum_{i \in I} x_i^n = 1 \quad \forall\, n \in N \tag{5.32}$$

$$x_i^n \geq 0 \quad \forall\, i \in I, n \in N \tag{5.33}$$

(5.30)–(5.33) is referred to as *m-SP*$_{Ret}$.

Note that the above model is nonlinear. We will formulate similar nonlinear models for different versions of the multi-period SP problem. These models can be linearized and in all our computations we use the linearized formulation. In this section, we present the nonlinear form as it has a simpler notation to follow.

Theorem 3:
Consider *m-SP*$_{Ret}$ with N decision nodes distributed among different periods. Assume that we solve *1-SP*$_{Ret}$ for each of the N nodes of the multi-period model by only taking the immediate branches of each node. The optimal decisions of *m-SP*$_{Ret}$ for the nodes of the final stage will be equivalent to the optimal decisions of *1-SP*$_{Ret}$ for those final stage nodes.

Proof:
The objective function of *m-SP*$_{Ret}$ can also be written as

$$z = \sum_{s \in S} \left(\prod_{b \in B_s} P_b . R_b \right) \tag{5.34}$$

where B_s is the set of branches on scenario s, $R_b = \sum_{i \in I} x_i^b \, (1 + \alpha_i + \beta_i . m_b)$ where x_i^b is the proportion of stock i in the portfolio of branch b and m_b is the market return on branch b, and P_b is the probability of branch b.

If we define
S' = set of partial scenarios when last stage branches are excluded from the scenarios of set S
$B_{S'}$= set of branches of partial scenario s'
$L_{S'}$= set of last stage branches of scenarios in S that are excluded to generate partial scenario s',
then z can also be written as

$$z = \sum_{s' \in S'} \left[\prod_{b' \in B_{s'}} P_{b'} . R_{b'} . \left(\sum_{l \in L_{s'}} P_l . R_l \right) \right] \tag{5.35}$$

$\sum_{l \in L_{s'}} P_l . R_l$ corresponds to the objective function of *1-SP$_{Ret}$* with scenario set $L_{S'}$.

$\prod_{b' \in B_{s'}} P_{b'} . R_{b'}$ and $\sum_{l \in L_{s'}} P_l . R_l$ are independent.

$\sum_{l \in L_{s'}} P_l . R_l$ for different $s' \in S'$ are independent.

Then, the optimal decisions of *m-SP$_{Ret}$* for the nodes of the final stage will be equal to the optimal decisions of *1-SP$_{Ret}$* for those final stage nodes. ■

We consider Theorem 3 unintuitive. Moreover, following the line of thought of its proof, we can also generalize this theorem to the case of K vs. L-period SP models of maximizing expected return where K > L. If we find the partial scenario trees of the K-period model that correspond to the last L periods of the horizon, and use these partial trees in L-period models, then the following will hold: the optimal decisions of the K-period model for the last L periods and the optimal decisions of the L-period models will be equivalent.

Remark 2:
Consider *m-SP$_{Ret}$* with N decision nodes distributed among different periods. Assume that we solve *1-SP$_{Ret}$* for each of the N nodes of the multi-period model by only taking the immediate branches of each node. The optimal decisions of *m-SP$_{Ret}$* and *1-SP$_{Ret}$* may not be equivalent for nodes of stages other than the final one.

Example 2:
Consider a two-period scenario tree with two branches for each node. Figure 34 shows the scenario tree where the probability of each branch appears above the branch and market returns for each branch are indicated in italic bold. Assume we have two stocks with $\bar{\alpha} = \bar{0}$, $\bar{\beta} = (0.8, 0.4)$.

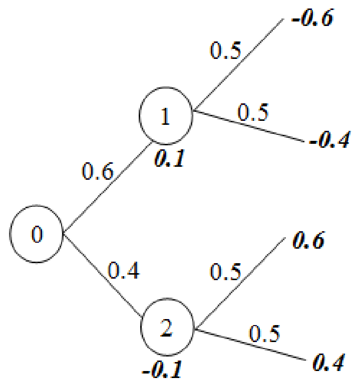


Figure 34. Scenario tree for Example 2

If we solve *1-SP_{Ret}* for nodes 0, 1 and 2, optimal decisions are realized as $\bar{x}(0) = (1, 0)$, $\bar{x}(1) = (0, 1)$, $\bar{x}(2) = (1, 0)$ where $\bar{x}(n)$ is the optimal stock weight vector for node *n*. On the other hand, the optimal decisions for *m-SP_{Ret}* are realized as $\bar{x}(0) = (\mathbf{0}, \mathbf{1})$, $\bar{x}(1) = (0, 1)$, $\bar{x}(2) = (1, 0)$ (The differences in the decisions of the multi-period models from those of the single-period models are indicated in bold throughout the examples).

The two-period expected return corresponding to the solutions of *1-SP_{Ret}* is 0.0336 whereas the optimal *m-SP_{Ret}* expected return is 0.0368. We see that *m-SP_{Ret}* has a better multi-period return value.

Theorem 3 is a general theorem and it is applicable to the TSM. As for Example 2 regarding Remark 2, we see that the positive first stage return is followed by negative second stage returns and vice versa. In our approach to the TSM, we employ the random walk model to generate the scenarios, and the random walk model allows this kind of pattern. So, decisions for stages other than the final one can differ for *m-SP_{Ret}* and *1-SP_{Ret}* in the TSM too.

Remark 3:
Consider *m-SP_{Ret}* with N decision nodes distributed among different periods. Assume that we solve *1-SP_{Ret}* for each of the N nodes of the multi-period model by only taking the immediate branches of each node. If we use the optimal solutions of *1-SP_{Ret}* models to calculate the corresponding multi-period return, the resulting value can never be better than the optimal *m-*

SP_{Ret} return. This directly follows from the use of the criterion of m-SP_{Ret} to assess the performance of 1-SP_{Ret} models. And we expect the optimal m-SP_{Ret} return to be better than the multi-period return of 1-SP_{Ret} models.

5.5.4 Minimizing CVaR in Single vs. Multi-period SP

Consider the multi-period SP model of minimizing CVaR:

$$min\ z = VaR_{\lambda} + \frac{1}{(1-\lambda)} \sum_{s \in S} p_s . a_s \tag{5.36}$$

s. t.

$$a_s \geq 1 - \prod_{n \in N_s} \left[\sum_{i \in I} x_i^n . (1 + \alpha_i + \beta_i . m_s^n) \right] - VaR_{\lambda} \quad \forall\, s \in S \tag{5.37}$$

$$\sum_{i \in I} x_i^n = 1 \quad \forall\, n \in N \tag{5.38}$$

$$a_s \geq 0 \quad \forall\, s \in S \tag{5.39}$$

$$x_i^n \geq 0 \quad \forall\, i \in I, n \in N \tag{5.40}$$

(5.36)–(5.40) is referred to as m-SP_{CVaR}.

Remark 4:

Consider m-SP_{CVaR} with N decision nodes distributed among different periods. Assume that we solve 1-SP_{CVaR} for each of the N nodes of the multi-period model by only taking the immediate branches of each node. The optimal decisions of m-SP_{CVaR} and 1-SP_{CVaR} models may not be equivalent; because 1-SP_{CVaR} only considers single-period scenarios whereas m-SP_{CVaR} considers a larger number of scenarios over a longer horizon. If we use the optimal solutions of 1-SP_{CVaR} models to calculate the corresponding multi-period CVaR, the resulting value can never be better than the optimal m-SP_{CVaR} CVaR. This directly follows from the use of the criterion of m-SP_{CVaR} to assess the performance of 1-SP_{CVaR} models. And we expect the optimal m-SP_{CVaR} CVaR to be better than the multi-period CvaR of 1-SP_{CVaR} models.

Example 3:

Consider a two-period scenario tree with four branches for each node. Figure 35 shows the scenario tree where market returns for each branch are illustrated. All branches have equal probability. Assume we have two stocks with $\bar{\alpha} = \bar{0}$, $\bar{\beta} = (0.8, 0.4)$.

If we solve 1-SP_{CVaR} with λ = 0.75 for nodes 0 to 4, optimal decisions are realized as $\bar{x}(0) =$ (0, 1), $\bar{x}(1) =$ (0, 1), $\bar{x}(2) =$ (0, 1), $\bar{x}(3) =$ (0, 1), $\bar{x}(4) =$ (0, 1) where $\bar{x}(n)$ is the optimal weight vector for node n. On the other hand, the optimal decisions for m-SP_{CVaR} with λ = 0.75 is realized as $\bar{x}(0) =$ (**0.23, 0.77**), $\bar{x}(1) =$ (0, 1), $\bar{x}(2) =$ (0, 1), $\bar{x}(3) =$ (0, 1), $\bar{x}(4) =$ (0, 1).

The two-period CVaR corresponding to the solutions of *1-SP*$_{CVaR}$ is 0.02846 whereas the optimal *m-SP*$_{CVaR}$ CVaR is 0.02811. We can see that the decisions of *m-SP*$_{CVaR}$ differ from *1-SP*$_{CVaR}$, and the difference is in favor of *m-SP*$_{CVaR}$ when the criterion is multi-period return.

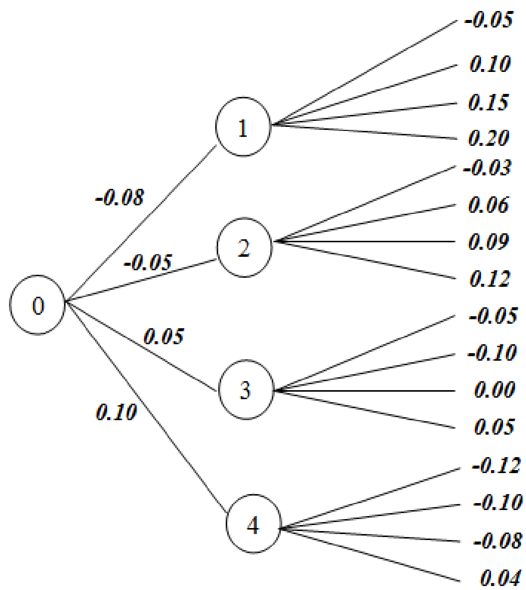


Figure 35. Scenario tree for Example 3

The scenario tree structure of Example 3 can be observed in the TSM when we use the random walk model to generate scenarios on market returns. Accordingly, we can expect the superiority of multi-period models over single-period models when the DM in the TSM has a multi-period planning horizon.

5.5.5 Minimizing CVaR - *ρ.Ret* in Multi-period SP

Consider the multi-period SP model of minimizing CVaR - *ρ.Ret*:

$$min\ z = VaR_\lambda + \frac{1}{(1-\lambda)}\sum_{s\in S} p_s.a_s - \rho.\sum_{s\in S} p_s.r_s \tag{5.41}$$

s. t.

$$r_s = \prod_{n\in N_s}\left[\sum_{i\in I} x_i^n.(1+\alpha_i+\beta_i.\mathrm{m}_s^{\mathrm{n}})\right] \quad \forall\, s\in S \tag{5.42}$$

$$a_s \geq 1 - r_s - VaR_\lambda \quad \forall\, s\in S \tag{5.43}$$

$$\sum_{i \in I} x_i^n = 1 \quad \forall\, n \in N \tag{5.44}$$

$$a_s \geq 0 \quad \forall\, s \in S \tag{5.45}$$

$$x_i^n \geq 0 \quad \forall\, i \in I, n \in N \tag{5.46}$$

(5.41)–(5.46) is referred to as *m-SP*$_{CVaR}$-*ρ.Ret*.

Remark 5:
The optimal solution of *m-SP*$_{CVaR}$-*ρ.Ret* and *m-SP*$_{CVaR}$ may be different. There can be alternative solutions that result in the same CVaR since CVaR only considers the worst (1- λ)% returns. Especially with high λ levels, there may be several solutions with the same CVaR.

Example 4:
Consider the scenario tree and stocks of Example 3. When minimizing CVaR with λ = 0.9375, we will only be considering the worst return over all scenarios. The two solutions below have the same optimal CVaR value of 0.05136.

Solution 1: $\bar{x}(0) = (0, 1)$, $\bar{x}(1) = (0, 1)$, $\bar{x}(2) = (1, 0)$, $\bar{x}(3) = (0, 1)$, $\bar{x}(4) = (0, 1)$ where $\bar{x}(n)$ is the optimal stock weight vector for node *n*.
Solution 2: $\bar{x}(0) = (0, 1)$, $\bar{x}(1) = (0, 1)$, $\bar{x}(2) = (0, 1)$, $\bar{x}(3) = (0, 1)$, $\bar{x}(4) = (0, 1)$.
Solution 1 has Ret = 0.0140 whereas solution 2 has Ret = 0.0080 and the two solutions differ only in their $\bar{x}(2)$ values.

In our applications with the TSM, we generally use λ = 0.9 and we often observe alternative solutions with the same CVaR.

5.5.6 Minimizing CVaR - *ρ.Ret* in Single vs. Multi-period SP

Consider the single-period SP model of minimizing CVaR- *ρ.Ret*:

$$min\ z = VaR_\lambda + \frac{1}{(1-\lambda)} \sum_{s \in S} p_s . a_s - \rho. \sum_{s \in S} p_s . r_s \tag{5.47}$$

s. t.

$$r_s = \sum_{i \in I} x_i \,.\, (1 + \alpha_i + \beta_i . m_s) \quad \forall\, s \in S \tag{5.48}$$

$$a_s \geq 1 - r_s - VaR_\lambda \quad \forall\, s \in S \tag{5.49}$$

$$\textstyle\sum_{i \in I} x_i = 1 \tag{5.50}$$

$$a_s \geq 0 \quad \forall\, s \in S \tag{5.51}$$

$$x_i \geq 0 \quad \forall\, i \in I \tag{5.52}$$

(5.47)–(5.52) is referred to as *1-*$SP_{CVaR\text{-}\rho.Ret}$.

Remark 6:
Consider *m-*$SP_{CVaR\text{-}\rho.Ret}$ with N decision nodes distributed among different periods. Assume that we solve *1-*$SP_{CVaR\text{-}\rho.Ret}$ for each of the N nodes of the multi-period model by only taking the immediate branches of each node. As we have discussed in Remark 4, if we use the optimal solutions of single-period models to calculate the corresponding multi-period CVaR, the resulting value can never be better than the optimal multi-period CVaR. Moreover, if we look at the multi-period return and CVaR values of the two models, there can be cases where *1-*$SP_{CVaR\text{-}\rho.Ret}$ results in dominated solutions.

Example 5:
Consider Example 3. The optimal solutions to *1-*SP_{CVaR} and *m-*SP_{CVaR} are also optimal for *1-*$SP_{CVaR\text{-}\rho.Ret}$ and *m-*$SP_{CVaR\text{-}\rho.Ret}$, respectively. The two-period CVaR and return corresponding to the solutions of *1-*$SP_{CVaR\text{-}\rho.Ret}$ are 0.02846 and 0.00825, respectively. The optimal CVaR and return values of *m-*$SP_{CVaR\text{-}\rho.Ret}$ are 0.02811 and 0.00863, respectively. We can see that the single period models result in a dominated solution.

5.5.7 Maximizing *Ret* - *ρ*. CVaR

We do not explicitly consider this case. In both single and multi-period models, having alternative solutions with the same maximum expected return is possible, but only under special combinations of α and β values of different stocks. In our applications in the TSM, we have not encountered such situations. There has always been a single stock which brings the maximum return value.

5.5.8 Maximizing *Ret* - *ρ*. CVaR subject to CVaR ≤ $\mathcal{E}$ in Single v.s Multi-period SP

Consider the single-period SP model of maximizing *Ret* - ρ. CVaR subject to CVaR ≤ Ɛ:

$$max\ z = \sum_{s \in S} p_s . r_s - \rho . \left(VaR_\lambda + \frac{1}{(1-\lambda)} \sum_{s \in S} p_s . a_s \right) \tag{5.53}$$

s. t.

$$r_s = \sum_{i \in I} x_i . (1 + \alpha_i + \beta_i . m_s) \quad \forall\, s \in S \tag{5.54}$$

$$a_s \geq 1 - r_s - VaR_\lambda \quad \forall\, s \in S \tag{5.55}$$

$$VaR_\lambda + \frac{1}{(1-\lambda)} \sum_{s \in S} p_s . a_s \leq \varepsilon \tag{5.56}$$

$$\sum_{i \in I} x_i = 1 \tag{5.57}$$

$a_s \geq 0 \quad \forall\, s \in S$ (5.58)

$x_i \geq 0 \quad \forall\, i \in I$ (5.59)

(5.53)–(5.59) is referred to as *1-SP*$_{Ret\ \text{-}\rho.CVaR}$.

Consider the multi-period SP model of maximizing *Ret* - ρ. CVaR subject to CVaR ≤ $\mathcal{E}$:

$$max\ z = \sum_{s \in S} p_s.r_s - \rho.\left(VaR_\lambda + \frac{1}{(1-\lambda)} \sum_{s \in S} p_s.a_s\right) \tag{5.60}$$

s. t.

$$r_s = \prod_{n \in N_s} \left[\sum_{i \in I} x_i^n.(1 + \alpha_i + \beta_i.m_s^n)\right] \quad \forall\, s \in S \tag{5.61}$$

$$a_s \geq 1 - r_s - VaR_\lambda \quad \forall\, s \in S \tag{5.62}$$

$$VaR_\lambda + \frac{1}{(1-\lambda)} \sum_{s \in S} p_s.a_s \leq \varepsilon \tag{5.63}$$

$$\sum_{i \in I} x_i^n = 1 \quad \forall\, n \in N \tag{5.64}$$

$$a_s \geq 0 \quad \forall\, s \in S \tag{5.65}$$

$$x_i^n \geq 0 \quad \forall\, i \in I, n \in N \tag{5.66}$$

(5.60)–(5.66) is referred to as *m-SP*$_{Ret\ \text{-}\rho.CVaR}$.

Remark 7:

Consider *m-SP*$_{Ret\ \text{-}\rho.CVaR}$ with N decision nodes distributed among different periods. Assume that we solve *1-SP*$_{Ret\ \text{-}\rho.CVaR}$ for each of the N nodes of the multi-period model by only taking the immediate branches of each node. If we use the optimal solutions of single-period models to calculate the corresponding multi-period return and CVaR, we may observe two cases that show that single-period models may not give optimal solutions for the multi-period problem:

- Let the corresponding multi-period CVaR of *1-SP*$_{Ret\ \text{-}\rho.CVaR}$ be R. R can be lower than the optimal CVaR of *m-SP*$_{Ret\ \text{-}\rho.CVaR}$. In this case the resulting multi-period return value of *1-SP*$_{Ret\ \text{-}\rho.CVaR}$ would be lower than the optimal *m-SP*$_{Ret\ \text{-}\rho.CVaR}$ return. If we solve *m-SP*$_{Ret\ \text{-}\rho.CVaR}$ with CVaR ≤ R, we may find the solution of *1-SP*$_{Ret\ \text{-}\rho.CVaR}$ to be dominated by that of *m-SP*$_{Ret\ \text{-}\rho.CVaR}$.
- The corresponding multi-period CVaR of *1-SP*$_{Ret\ \text{-}\rho.CVaR}$ can rise above $\mathcal{E}$. And it can be shown that the only condition that will allow the multi-period return of *1-SP*$_{Ret\ \text{-}\rho.CVaR}$ to be higher than that of *m-SP*$_{Ret\ \text{-}\rho.CVaR}$ is the infeasibility of *1-SP*$_{Ret\ \text{-}\rho.CVaR}$ by violating the CVaR ≤ $\mathcal{E}$ constraint.

Example 6:
Consider the scenario tree and stocks of Example 3. We take λ = 0.75 and $\mathcal{E}$ = 0.048. We choose this $\mathcal{E}$ value among the CVaR values feasible for $m\text{-}SP_{Ret\ \text{-}\rho.CVaR}$ and $1\text{-}SP_{Ret\ \text{-}\rho.CVaR}$ models for this example.

The optimal decisions of $1\text{-}SP_{Ret\ \text{-}\rho.CVaR}$ are realized as $\bar{x}(0) = (0.5, 0.5)$, $\bar{x}(1) = (1, 0)$, $\bar{x}(2) = (1, 0)$, $\bar{x}(3) = (0, 1)$, $\bar{x}(4) = (0, 1)$ where $\bar{x}(n)$ is the optimal stock weight vector for node n. On the other hand, the optimal decisions for $m\text{-}SP_{Ret\ \text{-}\rho.CVaR}$ are realized as $\bar{x}(0) = (\mathbf{1, 0})$, $\bar{x}(1) = (1, 0)$, $\bar{x}(2) = (1, 0)$, $\bar{x}(3) = (0, 1)$, $\bar{x}(4) = (0, 1)$.

The two-period return and CVaR values corresponding to the solutions of $1\text{-}SP_{Ret\ \text{-}\rho.CVaR}$ are 0.02422 and 0.03536, respectively. The optimal return and CVaR values of $m\text{-}SP_{Ret\ \text{-}\rho.CVaR}$ are 0.02462 and 0.04000, respectively. We can see that $1\text{-}SP_{Ret\ \text{-}\rho.CVaR}$ models resulted in a CVaR value lower than the CVaR of $m\text{-}SP_{Ret\ \text{-}\rho.CVaR}$, and thus failed to attain the optimal return value. Furthermore, if we solve $m\text{-}SP_{Ret\ \text{-}\rho.CVaR}$ with $\mathcal{E}$ = 0. 03536, the optimal return is realized as 0.02423, dominating the solution of $1\text{-}SP_{Ret\ \text{-}\rho.CVaR}$.

Example 7:
Consider a two-period scenario tree that we adopt from our applications with the TSM where the market returns are generated by the random walk model with the parameters of ISE-100. Figure 36 shows the scenario tree where market returns for each branch are illustrated. All branches have equal probability. Assume we have two stocks with $\bar{\alpha} = \bar{0}$, $\bar{\beta} = (0.8, 0.4)$.

With λ = 0.75 and $\mathcal{E}$ = 0.150, the optimal decisions of $1\text{-}SP_{Ret\ \text{-}\rho.CVaR}$ are realized as $\bar{x}(0) = (0.617, 0.383)$, $\bar{x}(1) = (0.617, 0.383)$, $\bar{x}(2) = (1, 0)$, $\bar{x}(3) = (0, 1)$, $\bar{x}(4) = (0, 1)$, $\bar{x}(5) = (0.831, 0.169)$ where $\bar{x}(n)$ is the optimal stock weight vector for node n. On the other hand, the optimal decisions for $m\text{-}SP_{Ret\ \text{-}\rho.CVaR}$ are realized as $\bar{x}(0) = (\mathbf{0.114, 0.896})$, $\bar{x}(1) = (\mathbf{1, 0})$, $\bar{x}(2) = (1, 0)$, $\bar{x}(3) = (0, 1)$, $\bar{x}(4) = (0, 1)$, $\bar{x}(5) = (\mathbf{0, 1})$.

The two-period return and CVaR values corresponding to the solutions of $1\text{-}SP_{Ret\ \text{-}\rho.CVaR}$ are 0.01305 and 0.185, respectively. The optimal return and CVaR values of $m\text{-}SP_{Ret\ \text{-}\rho.CVaR}$ are 0.01227 and 0.150, respectively. We can see that $1\text{-}SP_{Ret\ \text{-}\rho.CVaR}$ models resulted in an infeasible solution by violating the CVaR $\leq$ 0.150 constraint.

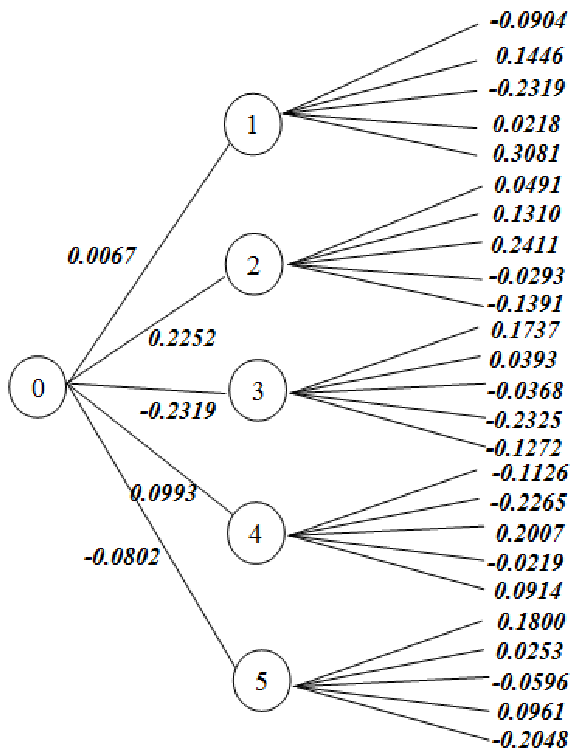


Figure 36. Scenario tree for Example 7

5.6 Rolling Horizon Approach

Our SP approaches provide the DM a set of decisions to implement throughout the planning horizon. To this point, we have treated these decisions as fixed, i.e., decisions of future periods are not updated after factors modeled by scenario trees are realized. Some researchers proposed to revise future decisions as time proceeds to reveal the up-to-date condition of markets. With rolling horizon approaches, multi-period models are used to make immediate decisions for the present. As time passes, new information becomes available and changes the future expectations. The models are then rolled so that the next decisions become immediate. Golub et al. (1995) proposed a two-stage rolling horizon SP model for fixed-income markets. Using mortgage securities, they implemented a dynamic strategy with the SP model and compared its performance to those of a single-period static model and a single-period stochastic model. The dynamic SP model was observed to produce superior results. Kouwenberg (2001) tested rolling horizon settings against a fixed-mix model for asset-liability management problems. Using a five-year planning horizon, they proposed to revise decisions at the beginning of each year by considering the years left of the horizon. They showed that SP models with rolling horizon settings perform particularly well with appropriately chosen scenario generation methods. Balıbek (2008) compared the performance of single and multi-period scenario trees that are updated after the immediate decisions in a multicriteria multi-period public debt management problem. Multi-period trees were observed to be superior for certain settings.

In this section, we consider bringing a rolling horizon approach to our SP applications and test its performance in comparison to the fixed scenario tree approach. Our scenario trees consist of branches that contain monthly ISE-100 returns generated by the random walk model (5.4). The mean and variance of ISE-100 return used in the random walk model are estimated by historical returns over a certain period. With the rolling horizon approach, we update the parameters of the random walk model every month to reveal the latest return. With the updated return mean and variance, we revise our scenario tree for the remaining months of the horizon. Figure 37 illustrates our approach on an example three-month setting with eight scenarios. We start with the three-month scenario tree (a). Let us assume that the index return represented by branch 1 is realized at the end of the first month. We update the parameters of the random walk model by adding this return to the set of historical returns. Then a new two-month scenario tree, (b), is generated. Let us assume now that branch 11ˈ is realized at the end of the second month. This return is also added to the set of historical returns and a final single-month tree (c) is generated from the updated random walk model. Finally, although it does not have an effect on scenario trees, let us assume that branch 111ˈˈ is realized for the third and final month. The optimal investment strategy for this setting is given by decisions of the first nodes of scenario trees (a), (b) and (c) for months one, two and three, respectively. The ending situation of the portfolio is given by the implementation of this strategy on the realized path 1-11ˈ-111ˈˈ.

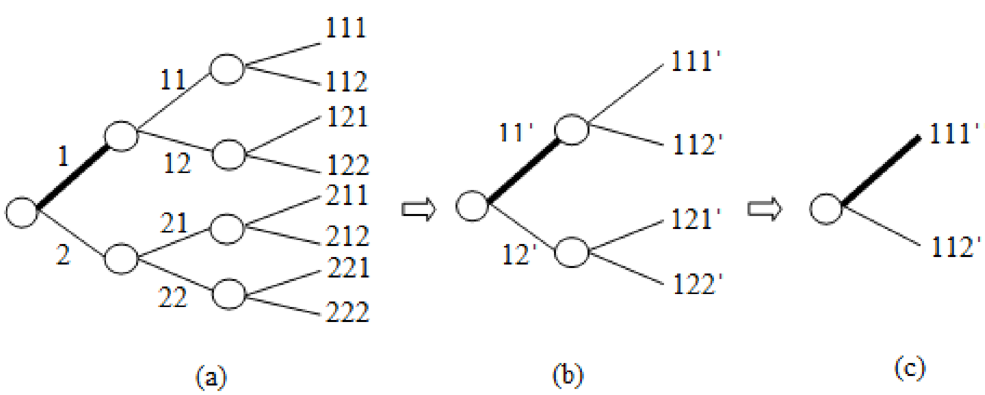


Figure 37. Rolling horizon scenario trees

Returns of stocks are estimated from the index return by the Single Index Model (5.3). As previously explained, the stock-specific return premiums and the sensitivities to the market are found by regression analysis using a set of historical index and stock returns. These intercept and coefficient values of regression may also be updated after a month of the horizon passes. However, we do not consider this update in our studies. One reason for this decision is the difficulty of finding an appropriate mechanism to simulate the stock return realizations. Since the index return is modeled by the random walk model, we can directly generate a random realization for it. On the other hand, stock returns are modeled by a regression equation. If we generate assumed realizations of stock returns by regressing them on the newly-generated index return, this will bring no difference to regression parameters. Thus, there will be no update to the regression equation. Using a totally different method to simulate stock returns does not seem appropriate as well, so we do not revise regression parameters. This decision may also be justified by the fact that one additional observation will in any case have minor effects considering the large set of past observations.

We do not expect the updated random walk model to bring substantial differences as well. Since our scenario generation technique is random, the only effect of the rolling horizon approach will be the minor update in the random walk parameters. Nevertheless, we make tests to compare the performance of rolling horizon against fixed scenario settings. To exemplify our comparison technique, we again refer to Figure 37. We have previously explained that the decisions of the rolling horizon setting will be given by the initial decisions of trees (a), (b) and (c). With the assumed scenario path, we can find the corresponding ending value. With the fixed-scenario tree, only tree (a) will be utilized, and decisions for three periods will be produced at time 0. To find the ending value of the fixed scenario tree, we will implement its decisions corresponding to the path most similar to the assumed realization. For example, we have previously assumed that the path 1-11ˈ-111ˈˈ is realized. After the initial decisions of time 0, we follow branch 1, and implement the corresponding decisions for the second period. After branch 1, now we continue with either 11 or 12, depending on how close they are in magnitude to 11ˈ. Let us assume that it is 11. Then the decisions of the ending node of branch 11 are finally implemented. These decisions for three periods are then applied on path 1-11ˈ-111ˈˈ to find the ending value of the portfolio.

We use the same 100 stocks used in Section 5.4 for our comparisons. We consider expected return and CVaR criteria, and exclude liquidity since it is independent of scenarios. Monthly return values corresponding to 2008-2009[10] are used to estimate parameters of the random walk and single index models. We avoid using a longer period of past returns since we want the update in the random walk model for the rolling horizon approach to be pronounced. The longer the period of past returns is, the smaller the effect of an additional point will be. We use 100 random instances to compare rolling and fixed scenario trees. In each instance, first we generate a 3-month scenario tree with five branches for each decision node, resulting in $5^3 = 125$ scenarios. This scenario tree is the initial tree of the rolling horizon approach and the sole tree for the fixed scenario tree approach. We randomly set a branch of the first period as the first realization, update the random walk model, and generate the second tree of two months with $5^2 = 25$ scenarios. From the first period of this tree, we again select a random branch, update the random walk model, and generate the final single-month tree of five scenarios. A final realization of index return is again set randomly on this tree.

[10] http://borsaistanbul.com/en/data/data/equity-market-data/equity-based-data

We use an augmented weighted Tchebycheff program to make portfolio decisions. The program is formulated as follows:

$$\text{Min } \alpha + \rho \sum_{i=1}^{k} (z_i^* - z_i) \tag{5.67}$$

$$\text{s.t.} \quad \alpha \geq w_i(z_i^* - z_i) \quad \forall\, i = 1, \dots, k \tag{5.68}$$

$$z_i = f_i(x) \qquad \forall\, i = 1, \dots, k \tag{5.69}$$

$$x \in X \tag{5.70}$$

where *z** is the ideal vector in the presence of k criteria, w_i is the weight of criterion *i* and ρ is a small positive number used as the augmentation factor. This program converges the efficient frontier at the solution determined by the weights chosen for criteria. We use expected return and CVaR at 90% probability as criteria. Normalization is applied to criteria values and *z** is taken as (1,1). For normalization, the distinct ideal and nadir points for each of the 100 instances are used. Weights of both criteria are taken as 0.5.

As previously stated, a random path of index return realizations is assumed for each of the 100 instances. As a result, for each instance, we obtain an ending portfolio return value for the rolling and fixed scenario tree settings. To compare the two settings with respect to expected return, we use the average of the ending return values of all instances. To make a comparison with respect to CVaR at 90% probability level, we use the average of 10 highest losses among these 100 instances. Table 8 shows the results.

Table 8. Comparing expected return and CVaR values of rolling and fixed scenario trees

	Expected Return (%)	CVaR (%)
Rolling Scenario Trees	24.539	-14.306
Fixed Scenario Trees	25.549	-14.212

We see that although the rolling horizon setting resulted in a better CVaR value, its expected return is lower than that of the fixed scenario tree setting. Studying the 100 instances individually, we observe that in 42 cases the rolling scenario trees resulted in higher expected return values than the fixed scenario trees, and in 44 cases the opposite situation holds. In the remaining 14 cases, the two settings are equivalent. These observations hint that there are no significant differences between the two approaches. To make a statistical test, we employ the paired *t*-test (see Hines et al., 2003, p. 292–293). Let D_j for j=1,2, …, 100 be the difference between the expected return values of rolling and fixed scenario trees in instance *j*. We calculate these values by subtracting the expected return of the fixed scenario tree from that of the rolling scenario tree. Our hypotheses are:

$H_0 : \mu_D = 0$
$H_A : \mu_D \neq 0$

where μ_D is the mean of the differences between expected return values of rolling and fixed scenario trees. The test statistic is given by $t = \overline{D}/\left(S_D/\sqrt{n}\right)$ where $\overline{D}$ and S_D are the sample mean and standard deviation of the differences, respectively; and n is the sample size. Our t statistic is calculated as -1.9019. As the critical t value with 95% confidence is 1.9842, we fail to reject the null hypothesis. This test also confirms that the rolling horizon settings do not bring significant differences in our SP approaches. This conclusion is not unexpected since we presumed that the addition of one month's return value to past observations will not change the random walk model parameters substantially. In addition, scenarios generated with the updated random walk parameters do not necessarily reflect those updates because of the error term. With our current SP approach to PO, the use of rolling horizon settings is therefore not justified.

The rolling horizon approach would be meaningful and essential if the DM changed her/his preferences over time. Currently, we are using the same weights of criteria for the whole planning period. It is possible for the DM to have different priorities for different periods based on her/his current financial position and obligations. Besides, the risk attitude of the DM may change after past achievements. For example, if she/he has earned a good return on the initial investment in the first two months, she/he may be more risk-averse in the third month to hold on to that return. On the contrary, if the first months of the period resulted in a loss, she/he may be more aggressive in the latter months to reverse the situation. That is, the risk attitude of the DM may depend on her/his current wealth. For such DM behavior, rolling horizon settings are required to update investment decisions accordingly. We consider this as an interesting field of future study.

CHAPTER 6

AN INTERACTIVE APPROACH TO STOCHASTIC PROGRAMMING-BASED PORTFOLIO OPTIMIZATION

With our basic SP approach covered in Section 5.2, we have provided the DM a discrete representation of the efficient frontier. The DM is expected to select an individual solution among the set of presented ones. This may be a difficult decision considering the large number of solutions, especially when the DM does not readily have clear-cut preferences. In this section, we present an interactive multistage approach to provide the DM with a single solution according to her/his preferences. DM preferences will be elicited by the help of comparison of a limited number of solutions in consecutive iterations. Our approach is based on the Interactive Weighted Tchebycheff Procedure of Steuer and Choo (see Steuer, 1986, p. 419-455). This procedure makes use of the augmented weighted Tchebycheff program which was formulated in Section 5.6 by (5.67)–(5.70). We choose to utilize this procedure since the augmented weighted Tchebycheff program can generate any efficient solution with appropriately chosen weights. Furthermore, the procedure extracts preferences from the DM with the help of simple comparisons without overwhelming her/him. The DM is not required to make distinctive statements of preference.

In essence, the Interactive Weighted Tchebycheff Procedure aims to converge the best solution by eliciting preference information from the DM in consecutive iterations and contracting the weight space accordingly. We explain the procedure further by discussing its steps briefly. First, an ideal criterion vector is computed and objectives are normalized. A number of random weight vectors are generated for the criteria (initially, they are freely chosen from the interval [0,1]), and they are reduced to a predetermined number of most-different weight vectors using a filtering approach. Using each of the resulting weight vectors, augmented weighted Tchebycheff programs are solved to determine solutions with minimum weighted Tchebycheff distances from the ideal point. The resulting solutions are filtered to a preset number of solutions and presented to the DM. After the DM chooses her/his most preferred solution, the weight vector corresponding to this solution is determined. Centered around this weight vector, new, narrowed down ranges for weights of criteria are calculated. Another set of random weight vectors that obey the new ranges are generated and filtered. Again solving augmented weighted Tchebycheff programs, a new set of solutions concentrated in a neighborhood of the previously selected solution is generated. These are filtered and presented to the DM for her/him to make a selection. This procedure is continued for a predetermined number of iterations unless the DM wants to stop prematurely.

We adopt this procedure to include the preferences of the DM to our SP approach. In our scenario trees, the utilization of larger number of scenarios will represent the market behavior better, but will result in an increase in complexity. Hence, we will first generate a large number of scenarios and then cluster them. As another issue, we also want to account for the stochasticity of our solutions due to the randomness involved in the scenario generation process as in Balibek and Köksalan (2012). For this purpose, instead of a single scenario tree, we will work with multiple scenario trees and generate confidence ellipsoids around solutions. The DM will be asked to choose between these ellipsoids. Another modification we make to the approach is to carry the preferred solution (ellipsoid) of one iteration to the next. In the original method, in each iteration, new solutions concentrated around the previously selected solution are generated. However, due to the randomness involved in weight generation, there is no guarantee that we will proceed to a better solution in each iteration. The solution of the previous iteration may have better utility value for the DM than the newly generated solutions. As a remedy, in each iteration, we also present the DM the best solution obtained so far.

We provide a flowchart of our approach in Figure 38. We refer to each stage in the flow chart by a stage number (given in boxes to the left of the stages). N is the number of scenarios used to construct ellipsoids. V is the number of weight vectors generated in each iteration later to be filtered. P corresponds to the sample size of ellipsoids to be presented to the DM. Let us further explain the mechanisms used in some of the steps. In step 1, we generate a large number of scenarios for better market representation purposes and then cluster them to decrease complexity. We employ the clustering algorithm explained in Section 5.3, which makes use of the K-means algorithm by MacQueen (1967) for this purpose. As a reminder, the K-means algorithm clusters data points in a predetermined number of disjoint classes. Each cluster has a centroid, and the algorithm aims at minimizing the sum of distances of all points to the centroid.

In step 4, we use a filtering mechanism to achieve the most-different weight vectors. The method we use is the Method of First Point Outside the Neighborhoods (see Steuer, 1986, p. 314-318). In this method, the vectors to be filtered are listed randomly, and a distance metric and its threshold value, d, are selected. The first vector is retained by the filter as the seed vector. Examining the following vectors, the first vector in the list that is at least d units away from the seed is retained. Continuing with the other vectors, now the first vector that is at least d units away from both the seed and the next vector selected are retained. The procedure is continued in this manner, each time retaining the first vector that is significantly different from all the previously selected vectors, until the list is exhausted. Different levels of d are tried until the desired number of filtered vectors is achieved. This method is also used in step 7, where we filter ellipsoids to determine the ones to present to the DM. The centroids of ellipsoids are utilized for this purpose. At the first iteration, $2P$ ellipsoids found with the weights of the iteration are filtered down to P. At other iterations, the selected ellipsoid of the previous iteration is added to the top of the list of $2P$ ellipsoids. Due to the working mechanism of the filtering method, the first ellipsoid is always retained. As a result, we ensure that the best ellipsoid so far is not lost and it is presented to the DM at each iteration.

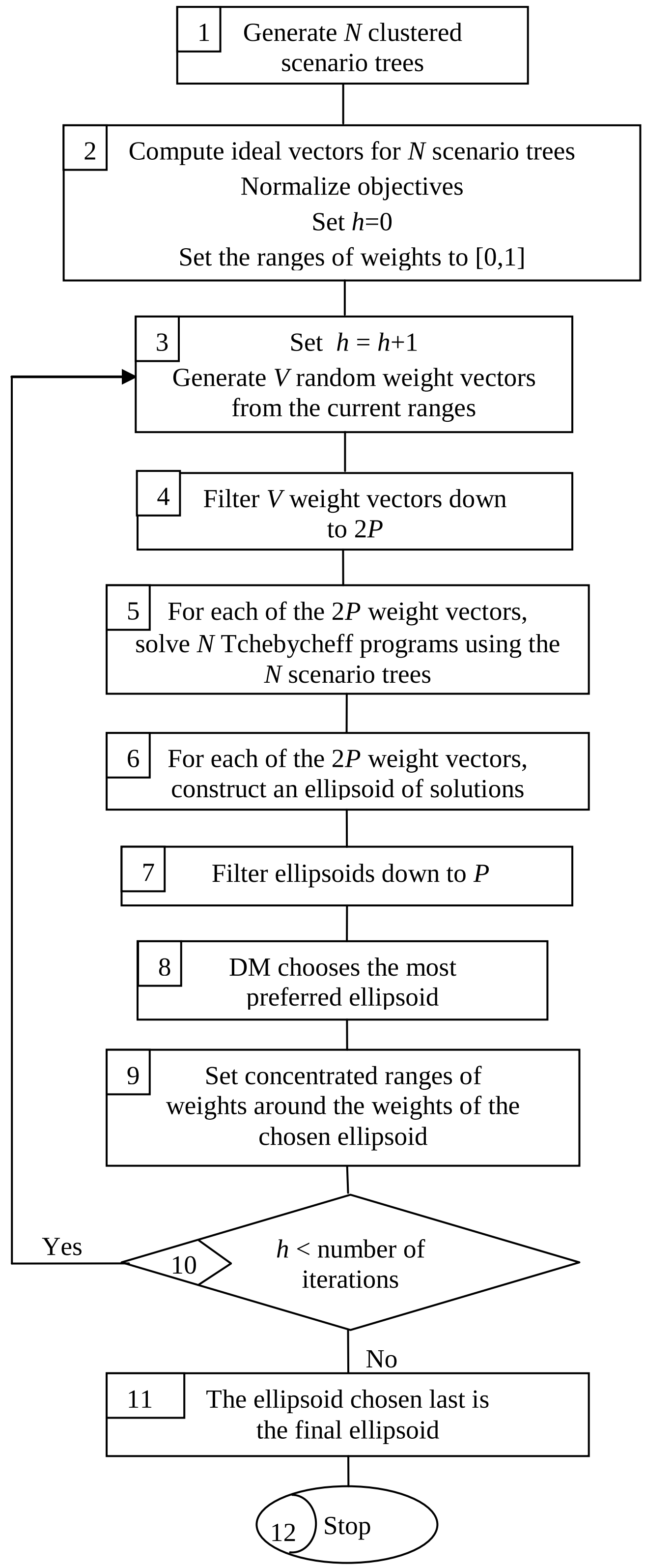


Figure 38. Flowchart of the Interactive Approach to SP-based PO

In step 5, for every weight vector filtered, we solve augmented weighted Tchebycheff programs and obtain N solutions for each. Utilizing the approach of Balibek and Köksalan (2012), we use these solutions to construct ellipsoids of solutions for them in step 6. (The reader is referred to Johnson and Wichern (2002, p. 210-238) for the theory on constructing confidence regions for multivariate means.) When the sample size is large, inferences about a population mean vector can be made without the normality assumption. Large-sample inferences about a mean vector are based on the χ^2 distribution. Let X_j's (j=1,…,n) be k-dimensional vectors sampled from a population with mean μ and covariance matrix S. When n-k is large, $n(\bar{X}-\mu)'S^{-1}(\bar{X}-\mu)$ is approximately χ^2 distributed with k degrees of freedom. Thus,

$$P\left[n(\bar{X}-\mu)'S^{-1}(\bar{X}-\mu) \leq \chi_k^2(\gamma)\right] = 1-\gamma \tag{6.1}$$

The above equality defines an ellipsoid which gives a 100(1-γ)% confidence region for the mean of X_j's. If we want to build simultaneous confidence intervals for the individual component means, we can project the ellipsoid on the axes of each component. This gives us the following 100(1-γ)% simultaneous confidence intervals:

$$\bar{x}_i \mp \sqrt{\chi_k^2(\gamma)\frac{s_{ii}}{n}} \quad i = 1, \dots, k \tag{6.2}$$

where s_{ii} is the variance for component i.

As mentioned before, centroids of the constructed 2P ellipsoids are used to filter them down to P in step 7.

For the DM to choose her/his most preferred ellipsoid in step 8, confidence intervals of all criteria corresponding to each ellipsoid are presented to her/him. This approach is preferred because ellipsoids are difficult to present and also hard to visualize for the DM with more than two criteria.

After the DM selects an ellipsoid, in step 9, new ranges for weights of criteria are generated. We require that these are concentrated around the weight vector corresponding to the solution in the ellipsoid that is closest to the centroid. We use unweighted Euclidean distance and normalized values to determine this solution, and refer to it as the 'pseudo centroid'. We apply the procedures defined by Steuer and Choo at this stage. Let z^h be the pseudo centroid of the chosen ellipsoid of iteration h. Then the most appropriate weight vector corresponding to z^h, denoted by w^h, is found as follows (see Steuer, 1986, p. 448):

$$
w_i^h = \begin{cases} \dfrac{1}{(z_i^* - z_i^h)} \left[\displaystyle\sum_{i=1}^{k} \dfrac{1}{(z_i^* - z_i^h)} \right]^{-1} & \text{if } z_i^h \neq z_i^* \text{ for all } i \\ 1 & \text{if } z_i^h = z_i^* \\ 0 & \text{if } z_i^h \neq z_i^* \text{ but } \exists j \ni \; z_j^h = z_j^* \end{cases} \tag{6.3}
$$

Next, the ranges of weight that are concentrated around w^h are formed as follows (see Steuer, 1986, p. 449):

$$
[LB_i^{h+1}, UB_i^{h+1}] = \begin{cases} [0, r^h] & \text{if } w_i^h - r^h/2 \leq 0 \\ [1 - r^h, 1] & \text{if } w_i^h + r^h/2 \geq 1 \\ [w_i^h - r^h/2,\; w_i^h + r^h/2] & \text{otherwise} \end{cases} \tag{6.4}
$$

where r is a pre-determined reduction factor for weights.

Below we provide a more formal description of the steps of our approach:

1. Generate N clustered scenario trees.
2. Compute ideal vectors for each scenario tree and normalize objectives.
   Set h=0, $[LB_i^{h+1}, UB_i^{h+1}] = [0\ ,\ 1]$ for all i.
3. h=h+1
   Generate V random weight vectors that obey LB_i^h and UB_i^h for all i; denote them as $\overline{rv}_j^h$, j = 1,…,V.
4. Filter the set of weight vectors $\overline{rv}_j^h$ down to a set of size 2P; denote the resulting vectors as $\overline{fv}_k^h$, k = 1,…,2P.
5. For each $\overline{fv}_k^h$, solve Tchebycheff programs with each of the N scenario trees. Let $s_{k,l}^h$ denote the solution found with weight vector $\overline{fv}_k^h$ and scenario tree l, l=1,…,N.
6. For each $\overline{fv}_k^h$, construct an ellipsoid of solutions using $s_{k,l}^h$ for l=1,…,N; denote them as e_k^h, k=1,…,2P.
7. Filter the set of ellipsoids e_k^h down to a set of size P, denote them as fe_m^h, m=1,…,P.
8. Among all fe_m^h, DM chooses an ellipsoid; let z^h denote the pseudo centroid of the chosen ellipsoid. Find the most appropriate weight vector corresponding to z^h with (6.3).
9. Find new ranges of weights, $[LB_i^{h+1}, UB_i^{h+1}]$, by utilizing (6.4).
10. If h is smaller than the desired number of iterations, stop. Otherwise, return to Step 3 and repeat.

We continue with applications of the approach with stocks from the ISE.

6.1 Results of Experiments

We use the same 100 stocks used in Section 5.4 for our experiments. Expected return, CVaR at 90% probability level and liquidity are used as criteria. The same set of historical data for the index and stock returns (2008-2009 monthly returns) and June 2011 data for liquidity are used. In Section 5.4, 2008-2009 monthly returns were preferred to amplify the effect of an additional observation in the rolling horizon scenario trees. In this section, we again avoid using longer periods that contain more distant return values that are less likely to represent the future state of the market.

We use a three-period SP setting with clustered scenario trees of 5^3=125 scenarios. The number of scenario trees to generate ellipsoids of solutions, N, is taken as 50. We use 90% confidence level for constructing ellipsoids. Ideal and nadir vectors of all scenario trees are computed individually. The nadir vectors are required for normalization purposes, and we approximate them by payoff nadirs. Normalized values are also used for filtering, and we use Euclidean distance as the distance metric of the filtering method.

Based on preliminary tests to determine good parameters for the approach, the number of iterations is set to 5. The sample size of solutions that are presented to the DM, P, is selected as 6. The number of weight vectors generated in each iteration, V, is chosen as 150, and 0.5 is used as r, the reduction factor for weights.

To make experiments with our approach, we need to assume an underlying preference function for the DM to simulate her/his selection of ellipsoids. With three types of distance metrics, we assume that the DM tries to minimize the weighted distance of normalized ellipsoid centroids from the ideal vector. These centroids are found by utilizing average normalized expected return, CVaR and liquidity values of the solutions in the ellipsoids. As distance metrics, we utilize Rectilinear, Euclidean and Tchebycheff distances. The weights of the DM used for expected return, CVaR and liquidity for all distances are 0.5, 0.25 and 0.25, respectively. We make three replications each for the three distance metrics. First, we present the details of one replication with the Tchebycheff distance to illustrate our approach. Five iterations of this replication are discussed. Later, we will summarize all of our results.

In applications of our approach with an actual DM, there is a possibility that she/he is inconsistent in her/his decisions. That is, her/his decision in an iteration may be in accordance with a different preference function than the other iterations. Even so, the approach provides flexibility to the DM to correct her/his direction of search in the latter iterations. After the DM chooses an ellipsoid, the new ranges of criteria weights are concentrated around the weights of this ellipsoid, but not very strictly. Unless the DM is exceedingly inconsistent, she/he will have room to adjust her/his preferences in the next iteration.

In normalization routines used in the weighted Tchebycheff programs and also the simulation of DM preferences of ellipsoids, we aim to scale criteria values between 0 and 1, where 0 is the worst and 1 is the best possible value for all criteria. However, because of the approximation of the nadir point by the payoff nadir, there may be cases where the lower bound is violated. Even so, the progress of the algorithm and the validity of results will not be affected. When simulating DM preferences, the ideal vector is taken as (1, 1, 1).

Expected return, CVaR and liquidity are denoted as *ret*, *cvar* and *liq* in vector representations. Let $\bar{v}_j^i = (v_{ret}, v_{cvar}, v_{liq})$ denote the weight vector numbered *j* (*j*=1, 2,…, 6) of iteration *i* that leads to one of the six ellipsoids in step 7 of the algorithm that are to be presented to the DM, where v_k denotes the weight of criterion *k*.

Iteration 1
150 random weight vectors are generated from the following ranges:

$[LB_{ret}^1, UB_{ret}^1] = [0, 1]$
$[LB_{cvar}^1, UB_{cvar}^1] = [0, 1]$
$[LB_{liq}^1, UB_{liq}^1] = [0, 1]$

150 weight vectors are filtered to 12. 50 augmented weighted Tchebycheff programs (one for each scenario tree) are solved with each of these 12 weight vectors. 12 ellipsoids of solutions are generated and filtered down to six by utilizing normalized ellipsoid centroids. For illustration purposes, Figure 39 shows the solutions used to construct the six ellipsoids. Ellipsoids are abbreviated as 'E' in illustrations. Six vectors of weights that led to these ellipsoids are:

$\bar{v}_1^1$ = (0.4145, 0.4065, 0.1790)
$\bar{v}_2^1$ = (0.0829, 0.5416, 0.3754)
$\bar{v}_3^1$ = (0.5798, 0.0853, 0.3348)
$\bar{v}_4^1$ = (0.7705, 0.2223, 0.0073)
$\bar{v}_5^1$ = (0.0690, 0.3625, 0.5684)
$\bar{v}_6^1$ = (0.2279, 0.1643, 0.6078)

For the six ellipsoids retained by the filter, the confidence regions of criteria are found by projecting the ellipsoids to individual criterion axes. Then the DM is asked to determine her/his most preferred one. Figure 40 illustrates the confidence regions of the six ellipsoids that are presented to the DM. The criteria values of the centroids of the ellipsoids are provided for the reader in Table 9, where we also include normalized values and the weighted Tchebycheff distance from the ideal vector.

Table 9. Criteria values of ellipsoids of iteration 1

				Normalized			
	Expected Return (%)	CVaR (%)	Liquidity	Expected Return	CVaR	Liquidity	Distance from Ideal Vector
E 1	20.9449	-16.3081	0.1965	0.6933	0.6873	0.2897	**0.1776**
E 2	17.6254	-15.3347	0.3165	0.4614	0.6424	0.4841	0.2693
E 3	20.2246	-5.0664	0.2530	0.6427	0.1671	0.3813	0.2082
E 4	24.6369	-19.4987	0.0157	0.9522	0.8342	-0.0033	0.2508
E 5	15.0347	-12.0301	0.4344	0.2808	0.4906	0.6752	0.3596
E 6	15.1378	-3.7627	0.4696	0.2857	0.1092	0.7322	0.3571

As observed from Table 9, Ellipsoid 1 has the smallest distance from the ideal vector. Accordingly, the DM is assumed to select Ellipsoid 1, and we continue to find the most appropriate weight vector corresponding to the pseudo centroid of this ellipsoid by using equation (6.3):

$w^1 = (w^1_{ret}, w^1_{cvar}, w^1_{liq})$ = (0.4145, 0.4065, 0.1790)

Centered around w^1, the new ranges of weights for the next iteration found by (6.4) are:

$[LB^2_{ret}, UB^2_{ret}] =$ [0.1645, 0.6645]
$[LB^2_{cvar}, UB^2_{cvar}] =$ [0.1565, 0.6565]
$\left[LB^2_{liq}, UB^2_{liq}\right] =$ [0, 0.5]

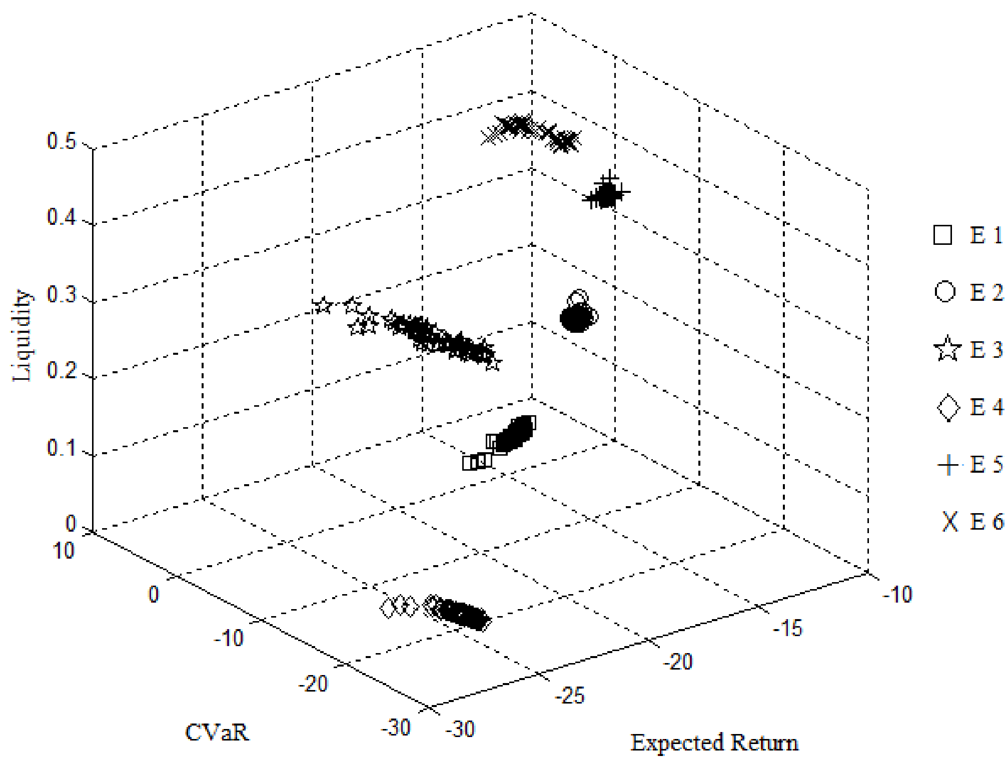


Figure 39. Plot of the solutions of the six ellipsoids of iteration 1

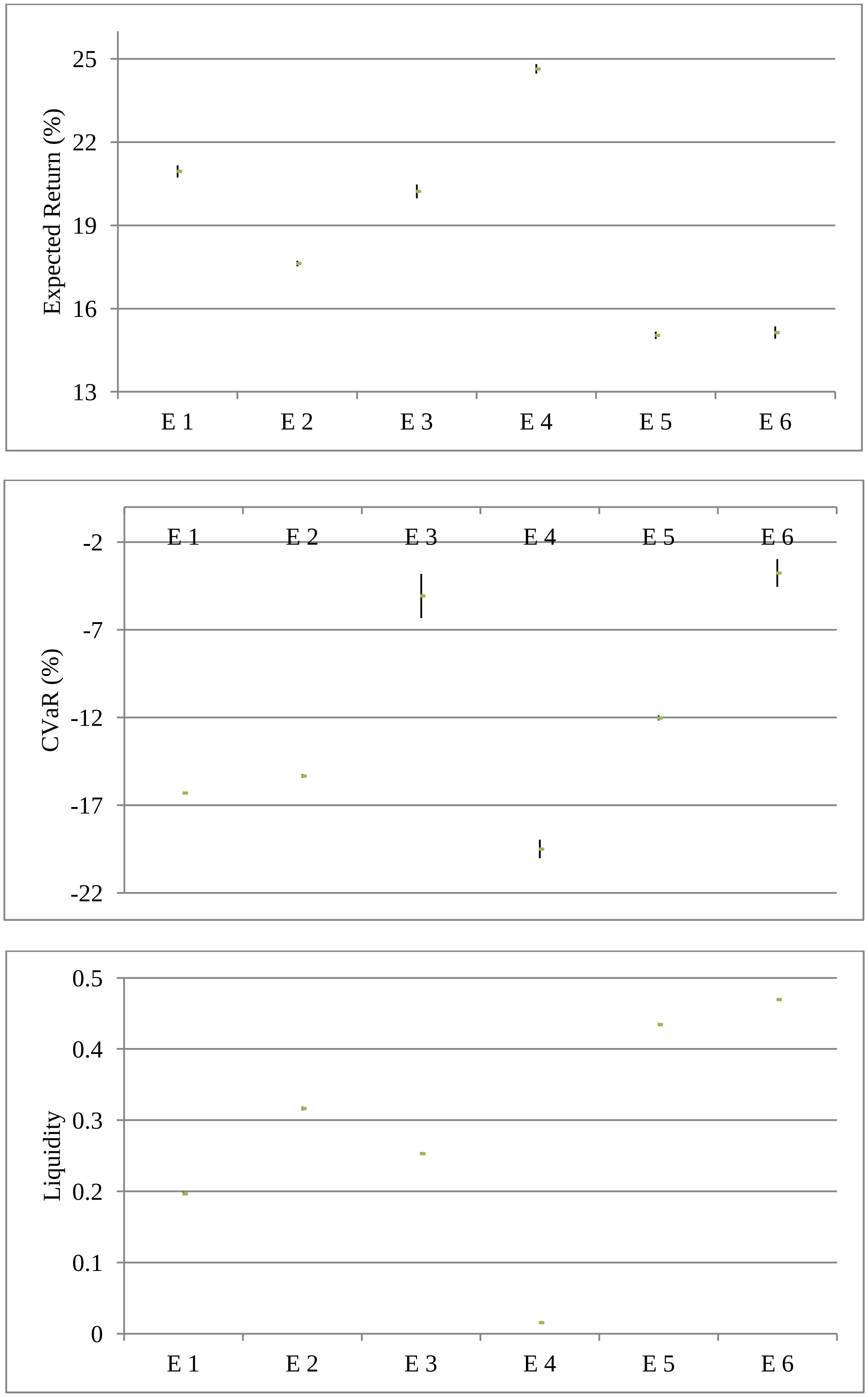


Figure 40. Presenting the DM ellipsoids of iteration 1

Iteration 2

150 random weight vectors from the current ranges are generated and filtered to most different 12. 50 augmented weighted Tchebycheff programs are solved for each, and 12 ellipsoids are formed. The preferred ellipsoid of the previous iteration is added to the top of the list of ellipsoids, and 13 ellipsoids are filtered to six. The corresponding weight vectors are :

$\bar{v}_1^2$ = (0.4145, 0.4065, 0.1790)
$\bar{v}_2^2$ = (0.1814, 0.4478, 0.3708)
$\bar{v}_3^2$ = (0.4146, 0.2039, 0.3815)
$\bar{v}_4^2$ = (0.2939, 0.2825, 0.4236)
$\bar{v}_5^2$ = (0.5441, 0.4117, 0.0442)
$\bar{v}_6^2$ = (0.4703, 0.2657, 0.2639)

Figure 41 illustrates the confidence regions of the six ellipsoids that are presented to the DM. The criteria values of the centroids of the ellipsoids and their distance from the ideal vector are provided in Table 10.

Table 10. Criteria values of ellipsoids of iteration 2

				Normalized			
	Expected Return (%)	CVaR (%)	Liquidity	Expected Return	CVaR	Liquidity	Distance from Ideal Vector
E 1	20.9449	-16.3081	0.1965	0.6933	0.6873	0.2897	0.1776
E 2	17.0272	-14.6160	0.3436	0.4194	0.6092	0.5280	0.2903
E 3	18.6685	-4.9015	0.3222	0.5337	0.1600	0.4933	0.2331
E 4	16.9977	-9.9364	0.3852	0.4168	0.3933	0.5953	0.2916
E 5	24.1288	-20.6807	0.0331	0.9161	0.8891	0.0251	0.2437
E 6	20.2814	-9.6149	0.2463	0.6468	0.3785	0.3704	**0.1766**

DM is assumed to select Ellipsoid 6 as it has the smallest distance from the ideal vector. We find the most appropriate weight vector corresponding to the pseudo centroid of Ellipsoid 6:

$w^2 = (w_{ret}^2, w_{cvar}^2, w_{liq}^2)$ = (0.4704, 0.2657, 0.2639)

Centered around w^2, the new ranges of weights for the next iteration are:

$[LB_{ret}^3, UB_{ret}^3]$ = [0.3454, 0.5954]
$[LB_{cvar}^3, UB_{cvar}^3]$ = [0.1407, 0.3907]
$[LB_{liq}^3, UB_{liq}^3]$ = [0.1389, 0.3889]

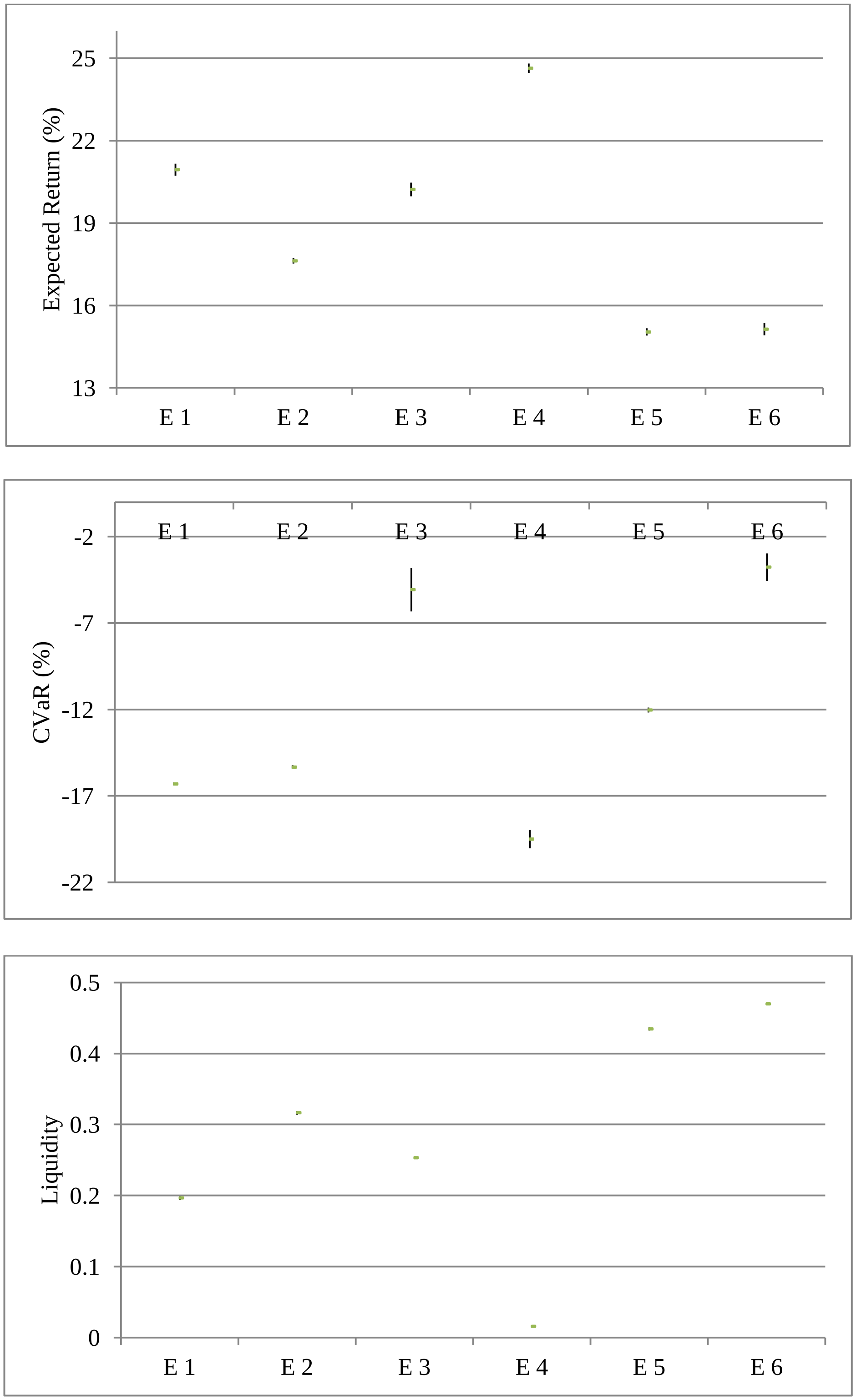


Figure 41. Presenting the DM ellipsoids of iteration 2

Iteration 3

The weight vectors of the six ellipsoids presented to the DM are :

$\bar{v}_1^3$ = (0.4703, 0.2657, 0.2639)
$\bar{v}_2^3$ = (0.3747, 0.3252, 0.3001)
$\bar{v}_3^3$ = (0.5603, 0.2639, 0.1758)
$\bar{v}_4^3$ = (0.4162, 0.2017, 0.3822)
$\bar{v}_5^3$ = (0.5052, 0.2031, 0.2917)
$\bar{v}_6^3$ = (0.3651, 0.2706, 0.3643)

Figure 42 illustrates the confidence regions of the six ellipsoids that are presented to the DM. The criteria values of the centroids of the ellipsoids and their distance from the ideal vector are provided in Table 11.

Table 11. Criteria values of ellipsoids of iteration 3

				Normalized			
	Expected Return (%)	CVaR (%)	Liquidity	Expected Return	CVaR	Liquidity	Distance from Ideal Vector
E 1	20.2814	-9.6149	0.2463	0.6468	0.3785	0.3704	**0.1766**
E 2	19.0309	-12.0786	0.2953	0.5592	0.4922	0.4497	0.2204
E 3	21.9319	-12.1670	0.1674	0.7624	0.4962	0.2425	0.1894
E 4	18.6768	-4.8334	0.3219	0.5343	0.1568	0.4928	0.2329
E 5	20.2210	-6.0574	0.2528	0.6425	0.2133	0.3808	0.1967
E 6	18.3354	-8.7695	0.3322	0.5105	0.3397	0.5095	0.2448

DM selects Ellipsoid 1 as it has the smallest distance from the ideal vector. Note that this is the ellipsoid carried from the previous iteration. The most appropriate weight vector corresponding to the pseudo centroid of Ellipsoid 1 was already found in iteration 2 as:

$w^3 = (w_{ret}^3, w_{cvar}^3, w_{liq}^3)$ = (0.4704, 0.2657, 0.2639)

Centered around w^3, the new ranges of weights for the next iteration are:

$[LB_{ret}^4, UB_{ret}^4] =$ [0.4079, 0.5329]
$[LB_{cvar}^4, UB_{cvar}^4] =$ [0.2032, 0.3282]
$[LB_{liq}^4, UB_{liq}^4] =$ [0.2014, 0.3264]

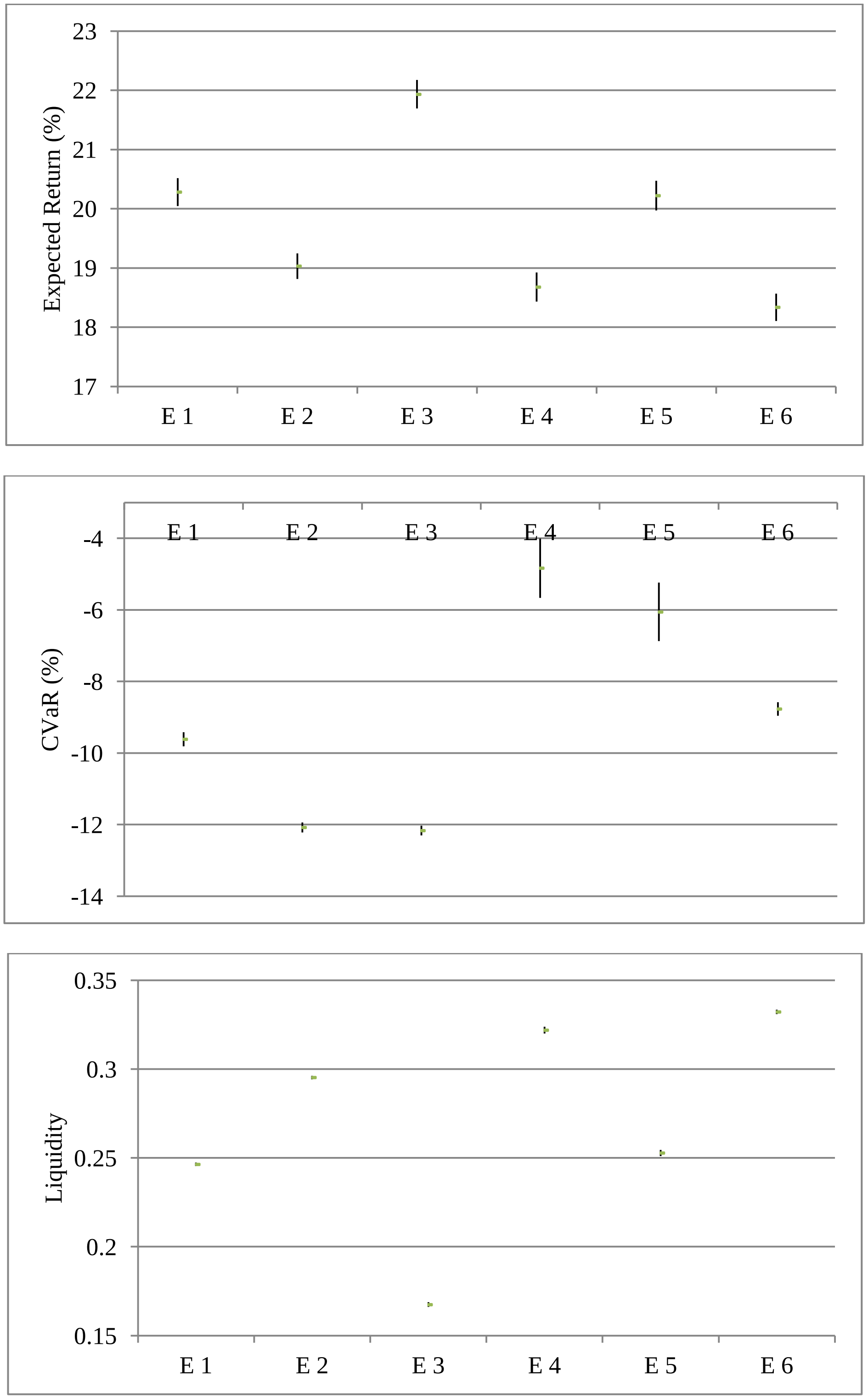


Figure 42. Presenting the DM ellipsoids of iteration 3

Iteration 4
The weight vectors of the six ellipsoids presented to the DM are :

$\bar{v}_1^4$ = (0.4703, 0.2657, 0.2639)
$\bar{v}_2^4$ = (0.4702, 0.2239, 0.3060)
$\bar{v}_3^4$ = (0.4866, 0.3086, 0.2048)
$\bar{v}_4^4$ = (0.4220, 0.2886, 0.2894)
$\bar{v}_5^4$ = (0.4816, 0.2810, 0.2373)
$\bar{v}_6^4$ = (0.4446, 0.3195, 0.2360)

Figure 43 illustrates the confidence regions of the six ellipsoids that are presented to the DM. The criteria values of the centroids of the ellipsoids and their distance from the ideal vector are provided in Table 12.

Table 12. Criteria values of ellipsoids of iteration 4

				Normalized			
	Expected Return (%)	CVaR (%)	Liquidity	Expected Return	CVaR	Liquidity	Distance from Ideal Vector
E 1	20.2814	-9.6149	0.2463	0.6468	0.3785	0.3704	0.1766
E 2	19.8287	-6.6630	0.2698	0.6150	0.2417	0.4084	0.1925
E 3	21.1128	-13.0062	0.2024	0.7050	0.5350	0.2992	0.1752
E 4	19.6085	-10.3990	0.2746	0.5996	0.4147	0.4162	0.2002
E 5	20.6403	-10.9932	0.2265	0.6741	0.4430	0.3386	**0.1654**
E 6	20.4242	-12.6929	0.2343	0.6556	0.5207	0.3511	0.1722

DM selects Ellipsoid 5 as it has the smallest distance from the ideal vector. The most appropriate weight vector corresponding to the pseudo centroid of Ellipsoid 5 is:

$w^4 = (w_{ret}^4, w_{cvar}^4, w_{liq}^4)$ = (0.4816, 0.2810, 0.2373)

Centered around w^4, the new ranges of weights for the next iteration are:

$[LB_{ret}^5, UB_{ret}^5]$ = [0.4504, 0.5129]
$[LB_{cvar}^5, UB_{cvar}^5]$ = [0.2498, 0.3123]
$[LB_{liq}^5, UB_{liq}^5]$ = [0.2061, 0.2686]

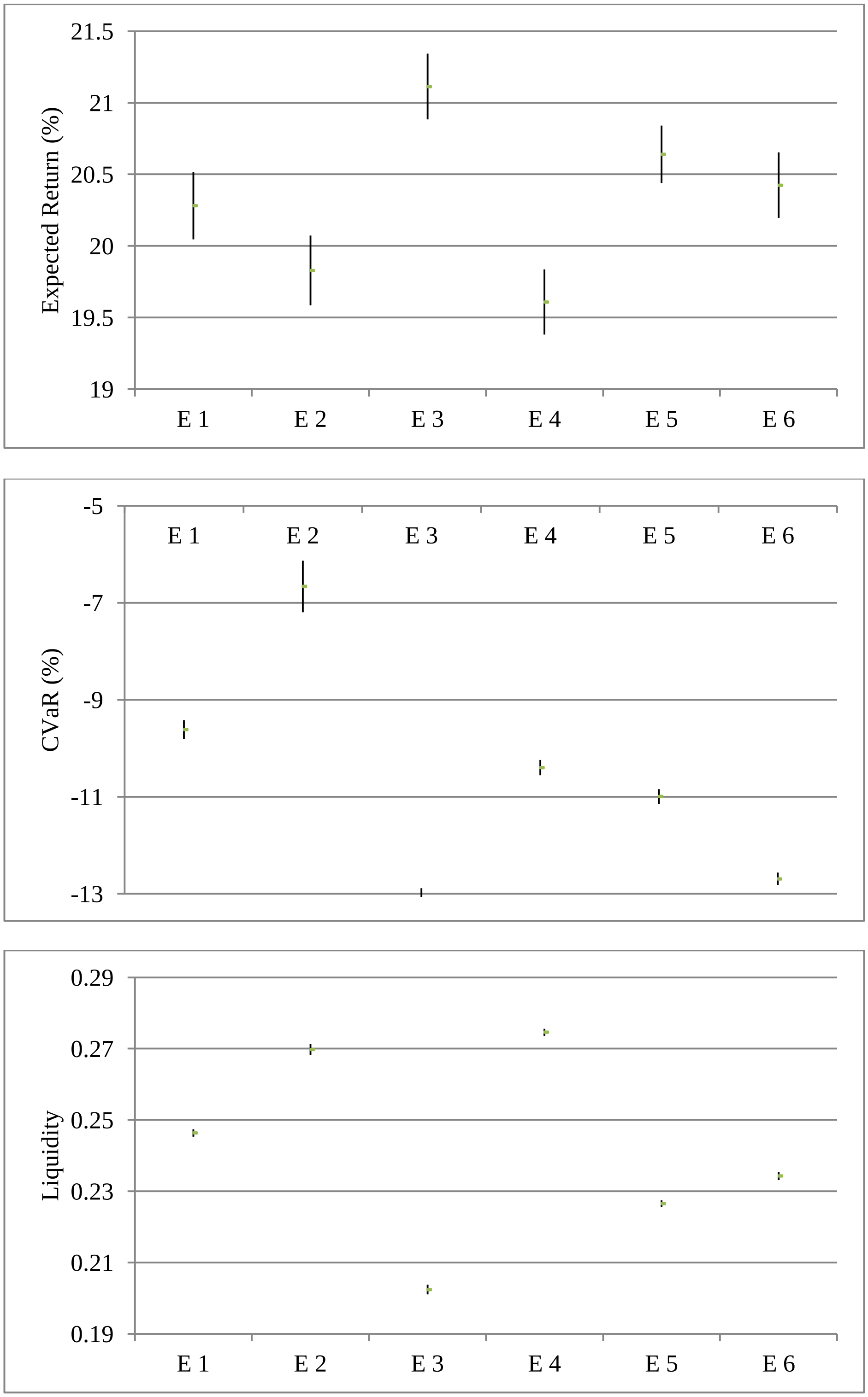


Figure 43. Presenting the DM ellipsoids of iteration 4

Iteration 5

The weight vectors of the six ellipsoids presented to the DM are :

$\bar{v}_1^5$ = (0.4816, 0.2810, 0.2373)
$\bar{v}_2^5$ = (0.4560, 0.2800, 0.2640)
$\bar{v}_3^5$ = (0.5124, 0.2559, 0.2317)
$\bar{v}_4^5$ = (0.4980, 0.2867, 0.2153)
$\bar{v}_5^5$ = (0.4784, 0.3117, 0.2099)
$\bar{v}_6^5$ = (0.4887, 0.2500, 0.2613)

Figure 44 illustrates the confidence regions of the six ellipsoids that are presented to the DM. The criteria values of the centroids of the ellipsoids and their distance from the ideal vector are provided in Table 13.

Table 13. Criteria values of ellipsoids of iteration 5

				Normalized			
	Expected Return (%)	CVaR (%)	Liquidity	Expected Return	CVaR	Liquidity	Distance from Ideal Vector
E 1	20.6403	-10.9932	0.2265	0.6741	0.4430	0.3386	**0.1654**
E 2	20.1703	-10.3560	0.2501	0.6390	0.4127	0.3765	0.1805
E 3	20.9464	-9.9543	0.2162	0.6933	0.3940	0.3217	0.1696
E 4	21.0546	-11.8585	0.2078	0.7009	0.4820	0.3081	0.1730
E 5	20.9903	-12.9868	0.2078	0.6965	0.5341	0.3081	0.1730
E 6	20.4440	-8.8494	0.2402	0.6581	0.3430	0.3604	0.1709

DM selects Ellipsoid 1 as it has the smallest distance from the ideal vector. Note that this is the ellipsoid carried from the previous iteration. The most appropriate weight vector corresponding to the pseudo centroid of Ellipsoid 1 was already found in iteration 4 as:

$w^5 = (w_{ret}^5, w_{cvar}^5, w_{liq}^5)$ = (0.4816, 0.2810, 0.2373)

If we were to make one more iteration, the new ranges of weights centered around w^5 would be:

$[LB_{ret}^6, UB_{ret}^6] =$ [0.4660, 0.4973]
$[LB_{cvar}^6, UB_{cvar}^6] =$ [0.2654, 0.2967]
$[LB_{liq}^6, UB_{liq}^6] =$ [0.2217, 0.2530]

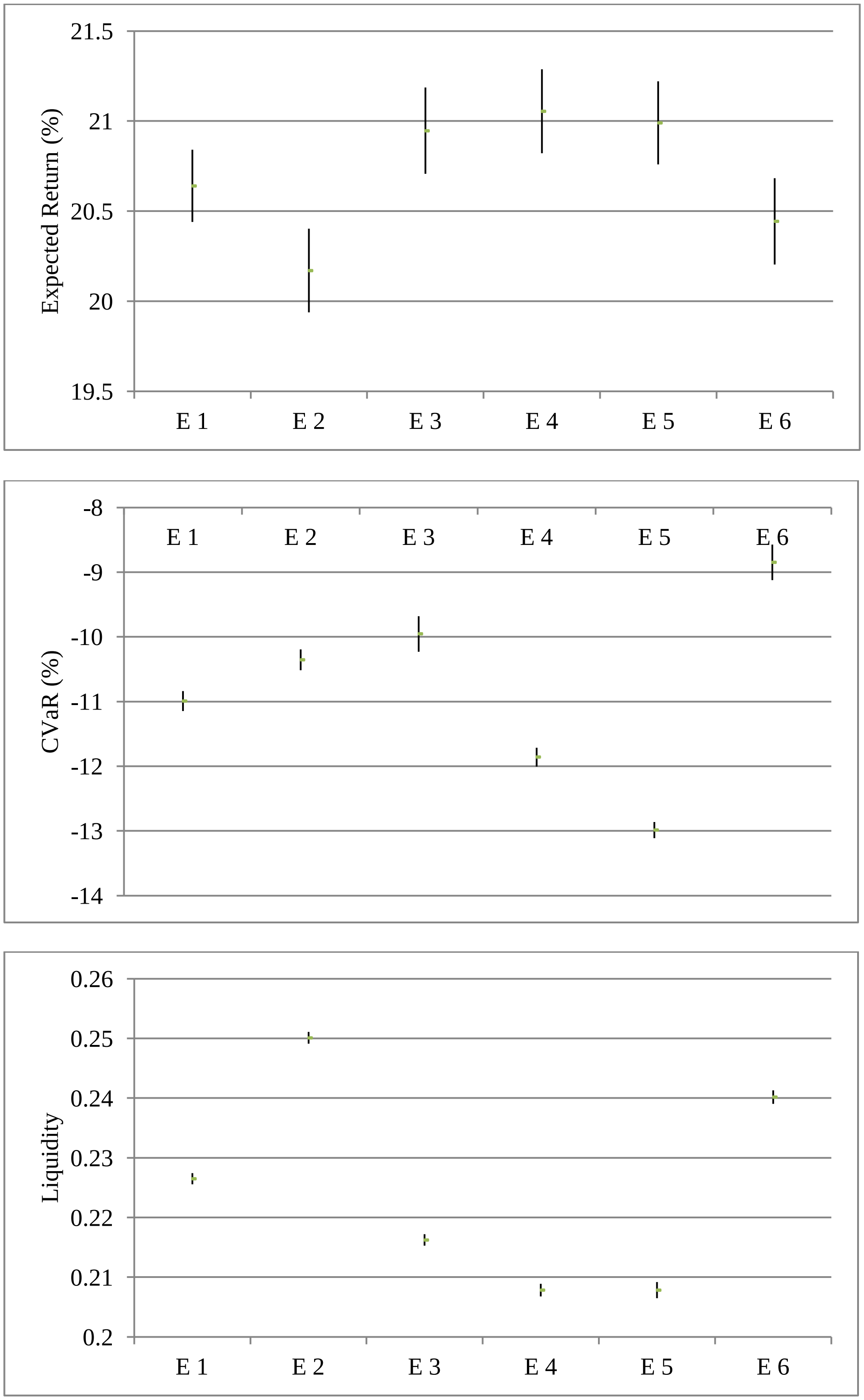


Figure 44. Presenting the DM ellipsoids of iteration 5

Table 14 summarizes the three replications performed while assuming the DM tries to minimize weighted Tchebycheff distance from the ideal point. The first replication consists of the five steps we just discussed. At the end of replications one, two and three, the DM is presented solutions with weighted Tchebycheff distances of 0.1654, 0.1698 and 0.1768 from the ideal point. To evaluate the results, we find the solution closest to the ideal point by using the DM's underlying weights of criteria, 0.5, 0.25 and 0.25 for expected return, CVaR and liquidity, respectively. This solution in normalized terms is (expected return, CVaR, liquidity) = (0.6725, 0.3561, 0.3450), and it has a distance value of 0.1637. We see that the algorithm produces solutions that are close to the best solution of the DM, and the achieved distance values are close to the minimum distance.

Tables 15 and 16 summarize the results with underlying Rectilinear and Euclidean distances for the DM. With Rectilinear distance, the best solution for the DM in normalized terms is (expected return, CVaR, liquidity) = (0.9723, 0.9895, -0.0068), and it has a distance value of 0.2907. With Euclidean distance, the solution is (0.7974, 0.8049, 0.1630) with a distance of 0.2375. We can observe that the algorithm can approach the best solution for the DM in these cases too. The distances of solutions attained while assuming a Euclidean distance for the DM are particularly good.

Table 14. Results of experiments with a Tchebycheff preference function for the DM

		Normalized Centroids			Corresponding Weights			Tchebychef Distance from Ideal Vector
		Expected Return	CVaR	Liquidity	Expected Return	CVaR	Liquidity	
Replication 1	Iter. 1	0.6933	0.6873	0.2897	0.4145	0.4065	0.1790	0.1776
	Iter. 2	0.6468	0.3785	0.3704	0.4704	0.2657	0.2639	0.1766
	Iter. 3	0.6468	0.3785	0.3704	0.4704	0.2657	0.2639	0.1766
	Iter. 4	0.6741	0.4430	0.3386	0.4816	0.2810	0.2373	0.1654
	Iter. 5	0.6741	0.4430	0.3386	0.4816	0.2810	0.2373	0.1654
Replication 2	Iter. 1	0.6822	0.1910	0.3397	0.5329	0.2107	0.2565	0.2022
	Iter. 2	0.6785	0.7107	0.2989	0.3892	0.4324	0.1784	0.1753
	Iter. 3	0.6785	0.7107	0.2989	0.3892	0.4324	0.1784	0.1753
	Iter. 4	0.6711	0.6336	0.3209	0.4199	0.3768	0.2033	0.1698
	Iter. 5	0.6711	0.6336	0.3209	0.4199	0.3768	0.2033	0.1698
Replication 3	Iter. 1	0.7509	0.5671	0.2482	0.5244	0.3018	0.1738	0.1879
	Iter. 2	0.7509	0.5671	0.2482	0.5244	0.3018	0.1738	0.1879
	Iter. 3	0.7256	0.3967	0.2878	0.5456	0.2441	0.2103	0.1780
	Iter. 4	0.7256	0.3967	0.2878	0.5456	0.2441	0.2103	0.1780
	Iter. 5	0.7027	0.2927	0.3159	0.5401	0.2252	0.2347	0.1768

Table 15. Results of experiments with a Rectilinear preference function for the DM

		Normalized Centroids			Corresponding Weights			Rectilinear Distance from Ideal Vector
		Expected Return	CVaR	Liquidity	Expected Return	CVaR	Liquidity	
Replication 1	Iter. 1	0.9522	0.8342	-0.0033	0.7489	0.2160	0.0351	0.3162
	Iter. 2	0.9522	0.8342	-0.0033	0.7489	0.2160	0.0351	0.3162
	Iter. 3	0.9522	0.8342	-0.0033	0.7489	0.2160	0.0351	0.3162
	Iter. 4	0.9522	0.8342	-0.0033	0.7489	0.2160	0.0351	0.3162
	Iter. 5	0.9522	0.8342	-0.0033	0.7489	0.2160	0.0351	0.3162
Replication 2	Iter. 1	0.9245	0.8284	0.0259	0.6603	0.2908	0.0490	0.3242
	Iter. 2	0.9245	0.8284	0.0259	0.6603	0.2908	0.0490	0.3242
	Iter. 3	0.9245	0.8284	0.0259	0.6603	0.2908	0.0490	0.3242
	Iter. 4	0.9451	0.8860	-0.0043	0.6510	0.3139	0.0351	0.3071
	Iter. 5	0.9451	0.8860	-0.0043	0.6510	0.3139	0.0351	0.3071
Replication 3	Iter. 1	0.8978	0.9296	0.0371	0.3918	0.5683	0.0399	0.3094
	Iter. 2	0.8978	0.9296	0.0371	0.3918	0.5683	0.0399	0.3094
	Iter. 3	0.8978	0.9296	0.0371	0.3918	0.5683	0.0399	0.3094
	Iter. 4	0.9347	0.9530	-0.0058	0.4077	0.5665	0.0259	0.2958
	Iter. 5	0.9347	0.9530	-0.0058	0.4077	0.5665	0.0259	0.2958

Table 16. Results of experiments with a Euclidean preference function for the DM

		Normalized Centroids			Corresponding Weights			Euclidean Distance from Ideal Vector
		Expected Return	CVaR	Liquidity	Expected Return	CVaR	Liquidity	
Replication 1	Iter. 1	0.6933	0.6873	0.2897	0.4145	0.4065	0.1790	0.2473
	Iter. 2	0.6933	0.6873	0.2897	0.4145	0.4065	0.1790	0.2473
	Iter. 3	0.7972	0.7110	0.1808	0.5130	0.3600	0.1270	0.2397
	Iter. 4	0.8341	0.7781	0.1311	0.5158	0.3857	0.0985	0.2391
	Iter. 5	0.7980	0.7341	0.1763	0.4987	0.3790	0.1223	0.2388
Replication 2	Iter. 1	0.9245	0.8284	0.0259	0.6603	0.2908	0.0490	0.2501
	Iter. 2	0.8890	0.6601	0.0883	0.6904	0.2255	0.0841	0.2495
	Iter. 3	0.8890	0.6601	0.0883	0.6904	0.2255	0.0841	0.2495
	Iter. 4	0.8890	0.6601	0.0883	0.6904	0.2255	0.0841	0.2495
	Iter. 5	0.9148	0.7719	0.0449	0.6844	0.2556	0.0600	0.2492
Replication 3	Iter. 1	0.8978	0.9296	0.0371	0.3918	0.5683	0.0399	0.2467
	Iter. 2	0.7940	0.7151	0.1835	0.5063	0.3660	0.1277	0.2395
	Iter. 3	0.7940	0.7151	0.1835	0.5063	0.3660	0.1277	0.2395
	Iter. 4	0.8150	0.7970	0.1478	0.4699	0.4281	0.1020	0.2377
	Iter. 5	0.7857	0.7773	0.1814	0.4496	0.4326	0.1177	0.2376

CHAPTER 7

CONCLUSIONS

PO is the problem of allocating funds between investment instruments in the financial market such as stocks, bonds, mutual funds, options and deposit accounts. The DM, personal or corporate, tries to maximize the ending wealth; and may also take other factors into account. This thesis addressed several approaches to multicriteria portfolio optimization.

We started with looking into the effects of different criteria on decision and objective spaces of the problem. For this purpose, we utilized single-period optimization settings with the criteria of expected return, variance, liquidity and CVaR. We also considered employing cardinality and weight constraints. We carried out computational studies using stocks from ISE and reported our findings. We first considered two-criteria models with all pairs of our four criteria. We provided insights into the conflict of criteria and the properties of the efficient frontiers of two-criteria models. We also looked into the effects of constraints on these issues. Next, we considered two cases of three-criteria models: expected return–variance–liquidity and expected return–variance–CVaR. We compared the results of two-criteria models to the results of the three-criteria models, both without and with constraints. We did not detect considerable improvements in values of the criterion that was previously disregarded. However, since the solutions of the two-criteria models were still inefficient to some extent for the three-criteria cases, we found it meaningful to take the third criterion into account even if only for tie-breaking purposes. Moreover, we showed that considering additional criteria results in enlarged regions in the efficient frontier that the DM may be interested in. The effects of constraints on the ranges of criteria in three-criteria models were also discussed. Lastly, we used our four criteria together and saw that adding the fourth criterion results in changes similar to the cases of adding the third criterion. There are no substantial improvements in the criterion that was previously omitted, but its range in the efficient frontier is enlarged. In conclusion, we showed that it is meaningful to consider multiple criteria in portfolio optimization. Recently, nonconventional investors with additional criteria to expected return and variance have been addressed in the literature; and we demonstrated the potential effects of these criteria. Different criteria result in different portfolios of investments, and by including additional criteria, the DM can be provided with different efficient regions that were previously unaccounted for. It is also possible to include constraints that limit the number of securities in a portfolio and the weight each security can take; and we illustrated the changes these constraints bring to investment options as well. As our insights on the effects of constraints on the criteria considered, expected return and liquidity are very sensitive to constraints that impose lower bounds on the number of stocks in the portfolio. Since these criteria attain their best values with a single stock, such constraints affect them greatly. Variance and CVaR, on the other hand, benefit from differentiation; and attain their best values with combinations of stocks. As a result, constraints that limit the maximum number of stocks were expected to have substantial effects on these criteria. However, variance and CVaR are observed to be quite flexible to such constraints; their best values attained with constraints are close the best values without constraints. Moreover, we can say that variance and CVaR are more flexible to changes in the available security pool than expected return and liquidity. These two risk measures can achieve good values with several different combinations of securities; but expected return and liquidity suffer greatly when securities that perform best with respect to them are lost.

We used genetic algorithm approaches as heuristics to solve PO problems with multiple criteria and complicating constraints. We applied a well-known genetic algorithm, NSGA-II, to solve PO with two and three criteria. Besides the conventional criteria of expected return and variance, we used a linear risk measure of expected returns below a DM-specified level. Using stocks from ISE, we compared the solutions of NSGA-II to exact efficient solutions and found that NSGA-II performs well with both two and three criteria. We also developed a genetic algorithm to solve expected return-variance PO with cardinality and weight constraints. We constrained the maximum number of securities that can be selected and the maximum weight each security can assume. These constraints complicated the formulation and introduced binary variables. The crossover and mutation operators were designed to obey the cardinality constraint and a repair mechanism was used to handle the weight constraint. We also included the DM in the process by conducting the search towards a reference point specified by her/him. Our computational studies showed that the proposed algorithm produces good convergence, especially when the number of securities is not large. Issues such as computational studies on the run time of the algorithm, extracting DM preferences by an improved procedure, being able to handle different types of constraints and testing different crossover operators were discussed as future studies. We consider genetic algorithms as useful tools for PO problems where there are many available securities, numerous criteria including nonlinear ones, and complicating constraints.

We also developed SP approaches to the multicriteria multi-period PO problem. We represented the uncertainties of the financial markets by discrete scenarios. Our scenario generation technique followed from our discussions on efficient market hypotheses and random walk models. After reviewing the literature on market efficiency and conducting statistical tests, we assumed that the TSM has weak-form market efficiency and thus follows a random walk model. We derived individual stock returns from a market-representative index using the Single Index Model. We employed expected return, CVaR and liquidity as our criteria in our model. Using stocks from ISE, we carried out computational studies, reported our findings and commented on our observations. This study is important in the sense that it is the first SP study that addresses multi-period multicriteria PO with CVaR. We also considered increasing the approximation capabilities of our scenario trees to represent the market behavior better. For this purpose, we generated a large number of scenarios, and then clustered them into classes of similarity to decrease computational complexity.

As another study in SP approaches to PO, we provided a detailed discussion on the properties of optimal SP solutions with different objectives and time periods. One goal was to see if single-period models can be used find optimal solutions to multi-period problems. This was considered a possibility since we do not consider assets with multi-period maturities and our scenario generation procedure involves randomness. First, with single-period models, we discussed the characteristics of optimal solutions when the objective is to maximize expected return or minimize CVaR. We showed that when certain conditions are met, optimal solutions can be found by observation. Next, we compared single and multi-period models of maximizing expected return and minimizing CVaR. We showed that single-period models may not find the optimal solutions to the multi-period problems. So, even though single-period models offer reduced computational complexity, multi-period models are superior when the DM has a multi-period planning horizon. Later, we also considered models that consider expected return and CVaR together. We demonstrated that when the problem is of multi-period nature, single-period models can result in suboptimal or infeasible solutions with a single criterion, and dominated solutions when we consider two criteria. We discussed our results in relation to the TSM whenever applicable, thus earned our findings practical value.

With our SP approaches to this point, we provided the DM a set of fixed decisions to implement throughout the planning horizon. Later we considered rolling horizon settings where future decisions are revised as time proceeds to reveal the up-to-date condition of the stock market. We updated the parameters of the random walk model as a period passes to reveal the latest return of the market-representative index. With updated parameters, we revised the scenario tree for the remaining periods of the horizon. We tested the performance of this approach to fixed scenario trees. As we had presumed, revised scenario trees did not bring substantial differences. This was attributed to the working mechanisms of the procedures used in scenario generation. On the other hand, we believe that rolling horizon settings in SP can be useful for DMs who are likely to change their preferences over time. Preferences of a DM can depend on past achievements and current wealth. In such cases, the DM may wish to make decisions for the future periods that are different than what she/he had made at the beginning of the planning horizon.

Lastly, to provide the DM a single solution according to her/his preferences rather than a representation of the whole efficient frontier, we considered an interactive approach. The developed approach was based on the Interactive Weighted Tchebycheff Procedure, which aims to converge the best solution by eliciting preference information from the DM in consecutive iterations and contracting the weight space accordingly. In our approach, we also took the stochasticity of market movements modeled by scenarios into account and provided the DM with statistical information. Working with multiple clustered scenarios, we constructed confidence ellipsoids around solutions and the DM was asked to make her/his preferences considering these ellipsoids. According to her/his selection, an increasingly concentrated set of ellipsoids were generated in each iteration. We modified the Interactive Weighted Tchebycheff Procedure by preserving the best solution generated so far throughout the process. We carried out experimental studies by simulating DM preferences with three types of underlying preference functions. Utilizing stocks from ISE, we showed that the proposed approach produces good results.

The insights we gained that are specific to the approaches studied are discussed in related parts of the thesis. But in general, we can say that the PO problem is an extensive area of research. There are many available investment options, several probable constraints and different possible planning horizons. Moreover, each investor is likely to consider different criteria and have different preferences. The PO model to be used should be constructed carefully taking all these factors into account since the results depend on them greatly. We consider it essential that the DM is involved in the process starting from the problem definition phase. She/he should also be provided guidance after efficient solutions with respect to the criteria considered are generated. There are typically many efficient solutions from many different regions of the solution space. The DM can make a better and more informed decision if she/he is enlightened about the interaction and conflict of criteria. If it is possible to extract her/his preferences before the efficient solutions are generated, interactive approaches may also be useful tools to guide the DM.

There are several issues that we consider as future studies. The modifications and improvements considered for the genetic algorithm we proposed are already discussed. For our SP approaches, we may consider the use of multi-index models to estimate stock returns. Currently, we are regressing the return of individual stocks on the return of an index. It is also possible to consider multiple factors that are likely to affect stock returns. As an example, the Three-Factor Model of Fama and French (see Bodie at el., 2009, p. 336) has become the dominating multi-factor model of stock returns in both academic and industry applications. In this model, in addition to the market index, market capitalization and book-to-market-ratio are also considered as factors. Scenario generation for three factors would be more complex and require different procedures. Furthermore, additional types of securities such as bonds or mutual funds may be considered. As another issue, we consider generalizing the results of our work on the properties of optimal SP solutions with different criteria and time periods. We looked at the solutions of single vs. multi-period models; now we can enhance our studies by generalizations to short vs. long-horizon SP. In addition, the rolling horizon approach may be used for investment settings where the DM changes preferences over time.

REFERENCES

Abdelaziz F.B., Aouni B. and Fayedh R.E. (2007), Multi-objective Stochastic Programming for Portfolio Selection, *European Journal of Operational Research*, 177, p. 1811-1823.

Balıbek E. (2008), Multi-objective Approaches to Public Debt Management, PhD Dissertation, Middle East Technical University.

Balibek E. and Köksalan M. (2010), A Multi-objective Multi-period Stochastic Programming Model for Public Debt Management, *European Journal of Operational Research*, 205(1), p. 205-217.

Balibek E. and Köksalan M. (2012), A Visual Interactive Approach for Scenario-based Stochastic Multi-objective Problems and an Application, *Journal of the Operational Research Society*, p. 1-15.

Ballestero E. and Romero C. (1996), Portfolio selection: A Compromise Programming Solution, *Journal of the Operational Research Society*, 47(11), p. 1377-1386.

Ballestro E. (2001), Stochastic Goal Programming: A Mean-variance Approach, *European Journal of Operational Research,* 131, p. 476-481.

Bana e Costa C.A. and Soares J.O. (2004), A Multicriteria Model for Portfolio Management, *European Journal of Finance*, 10(3), p. 198-211.

Basu S. (1983), The Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis, *Journal of Finance*, 32, p. 663-682.

Bertsimas D. and Shioda R. (2009), Algorithm for Cardinality-constrained Quadratic Optimization, *Computational Optimization and Applications*, 43, p. 1-22.

Bertsimas D., Darnell C. and Soucy R. (1999), Portfolio Construction through Mixed-integer Programming at Grantham, Mayo, Van Otterloo and Company, *Interfaces*, 29(1), p. 49-66.

Bodie Z., Kane A. and Marcus A.J. (2009), Investments, McGraw-Hill International Edition.

Buguk, C. and Brorsen, W. (2003), Testing Weak-form Market Efficiency: Evidence from the Istanbul Stock Exchange, *International Review of Financial Analysis*, 12, p. 579-590.

Campbell, J.Y., Lo, A.W. and MacKinlay, A.C. (1997), The Econometrics of Financial Markets, Princeton University Press.

Cesarone F., Scozzari A. and Tardella F. (2012), A New Method for Mean-variance Portfolio Optimization with Cardinality Constraints, *Annals of Operations Research*, DOI 10.1007/s10479-012-1165-7.

Chang T.-J., Meade N., Beasley J.E. and Sharaiha Y.M. (2000), Heuristics for Cardinality Constrained Portfolio Optimisation, *Computers & Operations Research*, 27, p. 1271-1302.

Conrad J. and Kaul G. (1988), Time-Variation in Expected Reyurns, *The Journal of Business*, 61(4), p. 409-425.

Deb K., Pratap A., Agarwal S. and Meyarivan T. (2002), A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II, *IEEE Transactions on Evolutionary Computation*, 6(2), p. 182-197.

Deb K. (2001), Multi-Objective Optimization using Evolutionary Algorithms, John Wiley and Sons, Inc.

Dupacova J. (1999), Portfolio Optimization via Stochastic Programming: Methods of Output Analysis, *Mathematical Methods of Operations Research*, 50, p. 245-270.

Dupacova J., Consigli G. and Wallace S.W. (2000), Scenarios for Multistage Stochastic Programs, *Annals of Operations Research*, 100, p. 25-53.

Ehrgott M. and Tenfelde-Podehl D. (2001), Computing Nadir Values in Three Objectives, *Lecture Notes in Economics and Mathematical Systems*, 507, p. 219-228.

Ehrgott M., Klamroth K. and Schwehm C. (2004), An MCDM Approach to Portfolio Optimization, *European Journal of Operational Research*, 155, p. 752–770.

Fama E.F. and French K..R. (1993), Common Risk Factors in the Returns on Stocks and Bonds, *Journal of Financial Economics*, 33, p. 3-56.

Fang Y., Xue R. and Wang S. (2009), A Portfolio Optimization Model with Fuzzy Liquidity Constraints, *2009 International Joint Conference on Computational Sciences and Optimization*, *IEEE*, p. 472-476.

Golub B., Holmer M., McKendall R., Pohlman L and Zenios S.A. (1995), A Stochastic Programming Model for Money Management, *European Journal of Operational Research*, 85, p. 282-296.

Guastaroba G., Mansini R. and Speranza M.G. (2009), On the Effectiveness of Scenario Generation Techniques in Single-period Portfolio Optimization, *European Journal of Operational Research*, 192, p. 500-511.

Gülpınar N., Rustem B. and Settergren R. (2003), Multistage Stochastic Mean-Variance Portfolio Analysis with Transaction Costs, *Innovations in Financial and Economic Networks*, 3, p. 46-63.

Haimes Y.Y., Lasdon L.S. and Wismer D.A. (1971), On a Bicriterion Formulation of the Problems of Integrated System Identification and System Optimization, *IEEE Transactions on Systems, Man, and Cybernetics*, 1(3), p. 296-297.

Hallerbach W. and Spronk J. (2002), The Relevance of MCDM for Financial Decisions, *Journal of Multicriteria Decision Analysis*, 11, p. 187-195.

Hines W.W., Montgomery D.C., Goldsman D.M. and Borror C.M. (2003), Probability and Statistics in Engineering, John Wiley and Sons, Inc.

Hirschberger M., Qi Y., Steuer R. E. (2007), Randomly Generating Portfolio-Selection Covariance Matrices with Specified Distributional Characteristics, *European Journal of Operational Research*, 177, p. 1610-1625.

Hoyland K. and Wallace S.W. (2001), Generating Scenario Trees for Multistage Decision Problems, *Management Science*, 47(2), p. 295-307.

Ibrahim K., Kamil A.A. and Mustafa A. (2008), Portfolio Selection Problem with Maximum Downside Deviation Measure: A Stochastic Programming Approach, *International Journal of Mathematical Models and Methods in Applied Sciences*, 1(2), p. 123-129.

Johnson R.A. and Wichern D.W. (2002), Applied Multivariate Statistical Analysis, Prentice Hall.

Kaut M. and Wallace S.W. (2003), Evaluation of Scenario-Generation Methods for Stochastic Programming, *Stochastic Programming E-Print Series*, 14, http://www.speps.org.

Kendall M. and Stuart A. (1979), The Advanced Theory of Statistics, Volume 2: Inference and Relationship, Charles Griffin & Company Limited.

Konno H. (1990), Piecewise Linear Risk Function and Portfolio Optimization, *Journal of Operational Research Society*, 33, p. 139-156.

Konno H. and Yamazaki H. (1991), Mean Absolute Deviation Portfolio Model and Its Applications to Tokyo Stock Market, *Management Science*, 37, p. 519-531.

Kouwenberg R. (2001), Scenario Generation and Stochastic Programming Models for Asset Liability Management, *European Journal of Operational Research*, 134, p. 279-292.

Lin C.-C. and Liu Y.-T. (2008), Genetic Algorithms for Portfolio Selection Problems with Minimum Transaction Lots, *European Journal of Operational Research*, 185, p. 393–404.

Lo A.W. and MacKinlay A.C. (1988), Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test, *Review of Financial Studies*, 1, p. 41-66.

MacQueen J.B. (1967), Some Methods for Classification and Analysis of Multivariate Observations, *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, University of California Press, 1, p. 281-297.

Mansini R., Ogryczak W. and Speranza M.G. (2003), On LP Solvable Models for Portfolio Selection, *Informatica*, 14, p. 37-62.

Mansini R., Ogryczak W. and Speranza M.G. (2007), Conditional Value at Risk and Related Linear Programming Models for Portfolio Optimization, *Annals of Operations Research*, 152, p. 227-256.

Markowitz H. (1959), Portfolio Selection: Efficient Diversification of Investments, John Wiley and Sons, Inc.

Michalowski W. and Ogryczak W. (2001), Extending the MAD Portfolio Optimization Model to Incorporate Downside Risk Aversion, *Naval Research Logistics*, 48, p. 186-200.

Odabasi, A., Aksu, C. and Akgiray, V. (2004), The Statistical Evolution of Prices on the Istanbul Stock Exchange, *The European Journal of Finance*, 10, p. 510-525.

Ogryczak W. and Ruszczynski A. (2002), Dual Stochastic Dominance and Quantile Risk Measures, *International Transactions in Operational Research*, 9, p. 661-680.

Pınar M.Ç. (2007), Robust Scenario Optimization Based on Downside-risk Measure for Multi-period Portfolio Selection, *OR Spectrum*, 209, p. 295-309.

Poterba J. and Summers L. (1988), Mean Reversion in Stock Prices: Evidence and Implications, *Journal of Financial Economics*, 22, p. 27-59.

Rockafellar R.T. and Uryasev S. (2000), Optimization of Conditional Value at Risk, *Journal of Risk*, 2(3), p. 21-41.

Rockafellar R.T. and Uryasev S. (2002), Conditional Value-at-Risk for General Loss Distributions, *Journal of Banking & Finance*, 26, p. 1443-1471.

Roman D., Darby-Dowman K.H. and Mitra, G. (2007), Mean-Risk Models Using Two Risk Measures: A Multi-Objective Approach, *Quantitative Finance*, 7 (4), p. 443-458.

Sarr A. and Lybek T. (2002), Measuring Liquidity in Financial Markets, IMF Working Paper, WP/02/232.

Seyhun H.N. (1986), Insiders' Profits, Costs of Trading and Market Efficiency, *Journal of Financial Economics*, 16(2), p. 189-212.

Shing C. and Nagasawa H. (1999), Interactive Decision System in Stochastic Multiobjective Portfolio Selection, *International Journal of Production Economics*, 60-61, p.187-193.

Skolpadungket P., Dahal K. and Harnpornchai N. (2007), Portfolio Optimization using Multi-objective Genetic Algorithms, *IEEE Congress on Evolutionary Computation 2007*, p. 516-523.

Smith, G. and Ryoo, H-J. (2003), Variance Ratio Tests of the Random Walk Hypothesis for European Emerging Stock Markets, *The European Journal of Finance*, 9, p. 290-300.

Steuer R.E. (1986), Multiple Criteria Optimization: Theory, Computation, and Application, Wiley International Edition.

Steuer R.E. and Na P. (2003), Multiple Criteria Decision Making Combined with Finance: A Categorized Bibliographic Study, *European Journal of Operational Research*, 150, p. 496-515.

Steuer, R.E., Qi Y. and Hirschberger M. (2006), Portfolio Optimization: New Capabilities and Future Methods, *Zeitschrift für Betriebswirtschaft*, 76(2), p. 199-219.

Steuer, R.E., Qi Y. and Hirschberger M. (2007a), Multiple Objectives in Portfolio Selection, *Journal of Financial Decision Making*, 1(1), p. 11–26.

Steuer, R.E., Qi Y. and Hirschberger M. (2007b), Suitable-Portfolio Investors, Nondominated Frontier Sensitivity, and the Effect on Standard Portfolio Selection, *Annals of Operations Research*, 152, p. 297-317.

Tuncer Şakar C. and Köksalan M. (2013a), Effects of Multiple Criteria on Portfolio Optimization, forthcoming in the *International Journal of Information Technology and Decision Making*.

Tuncer Şakar C. and Köksalan M. (2013b), A Stochastic Programming Approach to Multicriteria Portfolio Optimization, *Journal of Global Optimization*, 57(2), p. 299-314.

Xidonas P., Mavrotas G. and Psarras J. (2010), Equity Portfolio Construction and Selection Using Multiobjective Mathematical Programming, *Journal of Global Optimization*, 47, p. 185-209.

Xidonas P., Mavrotas G., Zopounidis C. and Psarras J. (2011), IPSSIS: An Integrated Multicriteria Decision Support System for Equity Portfolio Construction and Selection, *European Journal of Operational Research*, 210(2), p. 398-409.

Yu L., Wang S., Wu Y. and Lai K.K. (2004), A Dynamic Stochastic Programming Model for Bond Portfolio Management, *Lecture Notes in Computer Science*, 3039/2004, p. 876-883.

Yu L-Y., Ji X-D. and Wang S-Y. (2003), Stochastic Programming Models in Financial Optimization: A Survey, *Advanced Modeling and Optimization*, 5(1), p. 1-26.

Zopounidis C. (1999), Multicriteria Decision Aid in Financial Management, *European Journal of Operational Research*, 119, p. 404-415.

APPENDIX A

LISTS OF STOCKS USED FOR APPLICATIONS

Table 17. List of stocks used in Sections 3.1–3.3

ACIBADEM SAĞLIK	ATA YAT. ORT.
ADANA ÇİMENTO (A)	AKIN TEKSTİL
ADANA ÇİMENTO (B)	ATLAS YAT. ORT.
ADEL KALEMCİLİK	ATLANTİS YAT. ORT.
ADANA ÇİMENTO (C)	AVİVA SİGORTA
ANADOLU EFES	ALTINYUNUS ÇEŞME
AFM FİLM	AYEN ENERJİ
AFYON ÇİMENTO	AYGAZ
ATAKULE GMYO	BAGFAŞ
AKAL TEKSTİL	BAK AMBALAJ
AKBANK	BANVİT
AKÇANSA	BERDAN TEKSTİL
AK ENERJİ	BOSCH FREN SİSTEMLERİ
AKSİGORTA	BISAŞ TEKSTİL
AKSA	BEŞİKTAŞ FUTBOL YAT.
AKSU ENERJİ	BOLU ÇİMENTO
ALARKO HOLDİNG	BOSSA
ALARKO CARRIER	BOYNER MAĞAZACILIK
ALCATEL LUCENT TELETAŞ	BRİSA
ALARKO GMYO	BİRLİK MENSUCAT
ALKİM KAĞIT	BOROVA YAPI
ALKİM KİMYA	BORUSAN MANNESMANN
ALTERNATİFBANK	BORUSAN YAT. PAZ.
ALTINYILDIZ	BSH EV ALETLERİ
ALTINYAĞ	BATISÖKE ÇİMENTO
ANADOLU CAM	BATI ÇİMENTO
ANADOLU HAYAT EMEKLİLİK	BURSA ÇİMENTO
ANADOLU SİGORTA	BURÇELİK
ARÇELİK	BURÇELİK VANA
ARENA BİLGİSAYAR	ÇBS BOYA
ALTERNATİF YAT. ORT.	ÇELİK HALAT
ARSAN TEKSTİL	ÇEMTAŞ
ASELSAN	CEYLAN YATIRIM HOLDİNG
ASLAN ÇİMENTO	ÇİMSA
ANADOLU ISUZU	ÇELEBİ

Table 18. List of stocks used in Section 4.1.2 for the case of 10 stocks

ARÇELİK
BANVİT
DESA
DOĞAN YAYIN HOLDİNG
ECZACIBAŞI İLAÇ
GOODYEAR
IZOCAM
KENT GIDA
MİGROS
PETROL OFİSİ

Table 19. List of stocks used in Section 4.1.2 for the case of 100 stocks

ARÇELİK	BURSA ÇİMENTO	KRİSTAL KOLA
BANVİT	ÇBS BOYA	KÜTAHYA PORSELEN
DESA	COCA COLA İÇECEK	MEGES BOYA
DOĞAN YAYIN HOLDİNG	ÇBS PRINTAŞ	MENDERES TEKSTİL
ECZACIBAŞI İLAÇ	ÇİMSA	MERT GIDA
GOODYEAR	ÇELEBİ	MUTLU AKÜ
IZOCAM	DARDANEL	NUROL GMYO
KENT GIDA	DENİZBANK	PARSAN
MİGROS	DENTAŞ AMBALAJ	PINAR ET VE UN
PETROL OFİSİ	DERİMOD	RAY SİGORTA
ACIBADEM SAĞLIK	DOĞUŞ OTOMOTİV	SABANCI HOLDİNG
ADEL KALEMCİLİK	DYO BOYA	ŞİŞE CAM
ANADOLU EFES	EGE SERAMİK	ŞEKER PİLİÇ
AFM FİLM	EURO YAT. ORT.	TAÇ YATIRIM ORT
AK ENERJİ	EREĞLİ DEMİR ÇELİK	TAT KONSERVE
AKMERKEZ GMYO	ESCORT COMPUTER	TAV HAVALİMANLARI
ALARKO HOLDİNG	FENERBAHÇE SPORTİF	T.TUBORG
ALTINYILDIZ	FRİGO PAK GIDA	TURKCELL
ALTINYAĞ	FORD OTOSAN	TÜRK HAVA YOLLARI
ANADOLU CAM	GALATASARAY SPORTİF	TOFAŞ OTO FAB.
ARENA BİLGİSAYAR	GOLDAŞ KUYUMCULUK	TRAKYA CAM
ARSAN TEKSTİL	GRUNDİG ELEKTRONİK	TRABZONSPOR SPORTİF
ASELSAN	HÜRRİYET GAZETECİLİK	TÜRK TRAKTÖR
ANADOLU ISUZU	İDAŞ	TUKAŞ
ATLANTİS YAT. ORT.	İHLAS HOLDİNG	TÜPRAŞ
AVİVA SİGORTA	İNTEMA	ÜLKER BİSKÜVİ
AYGAZ	KAREL ELEKTRONİK	VAKKO TEKSTİL
BERDAN TEKSTİL	KARSAN OTOMOTİV	VANET
BOSCH FREN SİSTEMLERİ	KOÇ HOLDİNG	VESTEL
BİM MAĞAZALARI	KEREVİTAŞ GIDA	YATAŞ
BEŞİKTAŞ FUTBOL YATIRIM	KELEBEK MOBİLYA	YÜNSA
BSH EV ALETLERİ	KOZA DAVETİYE	ZORLU ENERJİ

Table 20. Expected returns of 25 stocks used in Section 4.2.2.1

Stock	Expected Return
1	-4.95
2	-5.87
3	-8.63
4	1.44
5	-7.97
6	0.55
7	3.94
8	-0.01
9	0.47
10	-1.89
11	-7.41
12	-0.22
13	4.80
14	8.69
15	-2.15
16	-8.14
17	0.83
18	1.88
19	-1.13
20	-5.74
21	10.12
22	-11.80
23	-2.55
24	-6.61
25	-3.18

Table 21. Expected returns of 50 stocks used in Section 4.2.2.1

Stock	Expected Return	Stock	Expected Return
1	-5.28	26	-5.92
2	3.03	27	-1.90
3	4.31	28	-1.65
4	-7.13	29	10.48
5	1.74	30	3.35
6	-2.28	31	-5.95
7	-4.68	32	-7.27
8	1.40	33	-3.61
9	-6.97	34	1.68
10	-6.91	35	9.33
11	-1.93	36	-3.11
12	-11.62	37	-0.90
13	-4.08	38	-7.97
14	-6.73	39	6.00
15	-8.95	40	2.93
16	-4.47	41	4.38
17	0.17	42	3.64
18	-1.61	43	-10.46
19	10.95	44	5.59
20	-0.90	45	-0.35
21	-4.49	46	-3.09
22	3.50	47	-2.32
23	3.05	48	5.42
24	-4.76	49	9.50
25	-0.63	50	5.07

Table 22. Covariance matrix of 25 stocks used in Section 4.2.2.1

σ_{ij}	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
1	71.20	-37.40	-6.71	0.79	16.37	-34.79	-58.32	-42.87	-19.67	25.96	30.13	7.65	27.21
2	-37.40	102.73	-12.03	23.33	22.56	9.68	121.41	87.40	45.59	70.69	-10.25	22.63	48.07
3	-6.71	-12.03	94.89	75.37	36.33	37.60	-9.28	88.24	14.01	-48.11	32.56	8.77	30.34
4	0.79	23.33	75.37	191.29	113.68	52.30	81.91	173.86	36.30	31.80	100.64	37.09	125.09
5	16.37	22.56	36.33	113.68	127.99	18.34	44.69	107.48	52.04	77.19	68.91	11.60	116.09
6	-34.79	9.68	37.60	52.30	18.34	153.37	51.72	61.22	17.92	17.89	55.41	-6.43	27.64
7	-58.32	121.41	-9.28	81.91	44.69	51.72	289.41	150.82	74.34	70.63	-15.79	-4.49	95.72
8	-42.87	87.40	88.24	173.86	107.48	61.22	150.82	243.91	32.67	34.30	58.00	20.97	140.12
9	-19.67	45.59	14.01	36.30	52.04	17.92	74.34	32.67	118.57	62.70	21.79	31.86	19.38
10	25.96	70.69	-48.11	31.80	77.19	17.89	70.63	34.30	62.70	184.65	70.03	32.08	78.98
11	30.13	-10.25	32.56	100.64	68.91	55.41	-15.79	58.00	21.79	70.03	114.66	37.24	61.58
12	7.65	22.63	8.77	37.09	11.60	-6.43	-4.49	20.97	31.86	32.08	37.24	47.56	9.59
13	27.21	48.07	30.34	125.09	116.09	27.64	95.72	140.12	19.38	78.98	61.58	9.59	156.62
14	-4.38	18.55	40.17	19.36	13.83	10.20	12.73	74.27	-38.55	-41.88	-13.36	-17.54	58.90
15	49.43	18.94	53.79	209.27	149.71	67.98	35.32	145.06	53.21	125.38	166.23	69.23	152.77
16	18.76	83.22	-9.83	90.96	72.65	12.77	168.78	110.62	32.29	88.89	9.55	14.32	131.44
17	22.12	78.39	59.05	85.09	110.90	21.06	37.02	133.30	55.62	88.99	43.22	35.44	133.87
18	15.25	-13.63	-26.70	53.46	20.27	65.17	-20.32	10.42	-44.42	50.92	69.85	11.19	34.24
19	-4.68	95.45	63.54	103.68	101.79	31.71	56.37	176.64	25.41	66.69	35.64	32.46	139.83
20	6.43	17.75	36.77	85.00	101.16	16.86	12.42	103.78	7.00	39.09	46.20	-3.91	104.27
21	57.31	38.62	42.76	115.03	123.22	-26.32	9.73	105.28	37.05	73.47	54.70	42.02	143.98
22	41.80	-14.63	33.42	123.82	96.20	57.75	118.12	83.48	38.10	76.13	83.77	-10.69	101.65
23	47.34	-31.23	29.88	41.51	99.70	43.13	-99.85	-19.05	74.79	105.07	93.54	28.65	31.21
24	35.69	37.11	-20.85	64.84	45.99	22.34	-34.46	72.52	-77.58	60.89	64.04	28.18	116.97
25	17.73	67.10	-15.47	108.61	48.48	-30.43	77.95	76.00	65.51	122.78	93.73	90.32	52.45

Table 22 (cont'd)

σ_{ij}	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**
1	-4.38	49.43	18.76	22.12	15.25	-4.68	6.43	57.31	41.80	47.34	35.69	17.73
2	18.55	18.94	83.22	78.39	-13.63	95.45	17.75	38.62	-14.63	-31.23	37.11	67.10
3	40.17	53.79	-9.83	59.05	-26.70	63.54	36.77	42.76	33.42	29.88	-20.85	-15.47
4	19.36	209.27	90.96	85.09	53.46	103.68	85.00	115.03	123.82	41.51	64.84	108.61
5	13.83	149.71	72.65	110.90	20.27	101.79	101.16	123.22	96.20	99.70	45.99	48.48
6	10.20	67.98	12.77	21.06	65.17	31.71	16.86	-26.32	57.75	43.13	22.34	-30.43
7	12.73	35.32	168.78	37.02	-20.32	56.37	12.42	9.73	118.12	-99.85	-34.46	77.95
8	74.27	145.06	110.62	133.30	10.42	176.64	103.78	105.28	83.48	-19.05	72.52	76.00
9	-38.55	53.21	32.29	55.62	-44.42	25.41	7.00	37.05	38.10	74.79	-77.58	65.51
10	-41.88	125.38	88.89	88.99	50.92	66.69	39.09	73.47	76.13	105.07	60.89	122.78
11	-13.36	166.23	9.55	43.22	69.85	35.64	46.20	54.70	83.77	93.54	64.04	93.73
12	-17.54	69.23	14.32	35.44	11.19	32.46	-3.91	42.02	-10.69	28.65	28.18	90.32
13	58.90	152.77	131.44	133.87	34.24	139.83	104.27	143.98	101.65	31.21	116.97	52.45
14	131.76	-37.60	-6.32	64.31	-40.09	87.82	64.69	31.10	-57.47	-88.97	104.37	-67.58
15	-37.60	328.67	132.10	136.04	127.64	130.61	91.45	184.61	172.09	170.35	116.81	177.38
16	-6.32	132.10	224.93	119.49	44.37	124.03	33.63	143.58	125.34	14.25	58.84	78.97
17	64.31	136.04	119.49	233.91	-8.82	226.71	102.38	198.23	19.96	109.24	109.18	33.47
18	-40.09	127.64	44.37	-8.82	146.49	11.67	14.08	22.40	52.44	56.41	112.06	46.57
19	87.82	130.61	124.03	226.71	11.67	250.65	112.12	186.09	-5.98	56.92	150.82	33.02
20	64.69	91.45	33.63	102.38	14.08	112.12	110.66	98.17	29.49	44.62	96.03	-4.42
21	31.10	184.61	143.58	198.23	22.40	186.09	98.17	231.34	58.05	114.53	118.35	68.81
22	-57.47	172.09	125.34	19.96	52.44	-5.98	29.49	58.05	265.19	90.71	-70.64	67.59
23	-88.97	170.35	14.25	109.24	56.41	56.92	44.62	114.53	90.71	280.96	-26.76	28.63
24	104.37	116.81	58.84	109.18	112.06	150.82	96.03	118.35	-70.64	-26.76	311.94	41.53
25	-67.58	177.38	78.97	33.47	46.57	33.02	-4.42	68.81	67.59	28.63	41.53	257.65

Table 23. Covariance matrix of 50 stocks used in Section 4.2.2.1

σ_{ij}	1	2	3	4	5	6	7	8	9	10	11	12	13
1	135.21	109.95	41.95	-14.42	5.59	83.25	21.85	48.52	107.02	99.80	52.59	78.90	92.69
2	109.95	267.71	44.49	69.99	18.13	170.58	98.10	148.31	140.46	141.72	35.64	97.85	84.51
3	41.95	44.49	43.44	15.31	9.86	35.72	25.88	37.99	57.79	43.62	52.77	51.51	40.96
4	-14.42	69.99	15.31	217.95	-38.42	94.93	-6.35	93.24	-22.82	84.32	51.71	65.61	-35.77
5	5.59	18.13	9.86	-38.42	68.65	-17.35	15.60	28.59	37.43	20.28	-8.69	-3.85	4.63
6	83.25	170.58	35.72	94.93	-17.35	182.10	66.03	133.37	83.07	98.67	38.31	44.28	25.45
7	21.85	98.10	25.88	-6.35	15.60	66.03	127.54	49.46	80.00	32.32	-7.55	45.22	17.65
8	48.52	148.31	37.99	93.24	28.59	133.37	49.46	152.41	69.80	120.64	40.23	39.08	-15.31
9	107.02	140.46	57.79	-22.82	37.43	83.07	80.00	69.80	155.71	67.98	37.35	71.37	71.31
10	99.80	141.72	43.62	84.32	20.28	98.67	32.32	120.64	67.98	199.18	73.54	123.51	39.00
11	52.59	35.64	52.77	51.71	-8.69	38.31	-7.55	40.23	37.35	73.54	100.90	85.51	80.92
12	78.90	97.85	51.51	65.61	-3.85	44.28	45.22	39.08	71.37	123.51	85.51	153.45	107.06
13	92.69	84.51	40.96	-35.77	4.63	25.45	17.65	-15.31	71.31	39.00	80.92	107.06	231.81
14	25.38	80.13	66.63	23.62	-35.74	63.89	96.31	50.15	83.24	36.58	80.64	80.84	32.38
15	102.31	219.51	72.56	12.10	24.56	146.76	115.81	103.81	186.11	39.14	57.80	76.51	139.69
16	89.42	117.37	2.01	-41.00	21.33	65.29	-16.31	44.72	41.78	66.32	3.74	10.46	97.75
17	97.56	114.12	47.26	63.56	0.76	79.29	-28.13	89.70	48.33	149.18	105.53	100.39	98.43
18	74.10	100.15	73.16	-10.39	33.90	57.03	68.57	65.60	123.51	65.34	84.60	87.63	99.62
19	68.50	148.14	45.98	74.84	60.32	72.16	51.05	113.99	108.72	134.83	52.88	103.15	45.05
20	52.37	137.19	21.24	15.26	47.16	92.33	26.01	102.79	97.15	38.20	-17.81	-21.91	-24.67
21	73.58	112.50	10.14	-7.73	-2.41	52.70	24.42	47.33	73.03	78.52	32.85	56.61	57.80
22	14.18	11.83	25.82	142.79	-28.36	55.82	-20.33	50.64	4.12	56.03	56.62	53.56	-29.78
23	71.17	82.76	33.61	82.08	-39.83	81.11	10.21	25.96	45.84	53.15	54.44	83.51	82.24
24	121.06	151.36	34.82	71.61	-14.69	93.90	70.40	54.63	91.32	158.47	47.89	158.75	90.72
25	41.38	102.52	18.75	14.27	8.25	8.10	8.10	24.51	30.33	83.84	43.02	89.56	100.46

Table 23 (cont'd)

σ_{ij}	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**
1	25.38	102.31	89.42	97.56	74.10	68.50	52.37	73.58	14.18	71.17	121.06	41.38
2	80.13	219.51	117.37	114.12	100.15	148.14	137.19	112.50	11.83	82.76	151.36	102.52
3	66.63	72.56	2.01	47.26	73.16	45.98	21.24	10.14	25.82	33.61	34.82	18.75
4	23.62	12.10	-41.00	63.56	-10.39	74.84	15.26	-7.73	142.79	82.08	71.61	14.27
5	-35.74	24.56	21.33	0.76	33.90	60.32	47.16	-2.41	-28.36	-39.83	-14.69	8.25
6	63.89	146.76	65.29	79.29	57.03	72.16	92.33	52.70	55.82	81.11	93.90	8.10
7	96.31	115.81	-16.31	-28.13	68.57	51.05	26.01	24.42	-20.33	10.21	70.40	8.10
8	50.15	103.81	44.72	89.70	65.60	113.99	102.79	47.33	50.64	25.96	54.63	24.51
9	83.24	186.11	41.78	48.33	123.51	108.72	97.15	73.03	4.12	45.84	91.32	30.33
10	36.58	39.14	66.32	149.18	65.34	134.83	38.20	78.52	56.03	53.15	158.47	83.84
11	80.64	57.80	3.74	105.53	84.60	52.88	-17.81	32.85	56.62	54.44	47.89	43.02
12	80.84	76.51	10.46	100.39	87.63	103.15	-21.91	56.61	53.56	83.51	158.75	89.56
13	32.38	139.69	97.75	98.43	99.62	45.05	-24.67	57.80	-29.78	82.24	90.72	100.46
14	251.78	145.51	-73.54	43.01	135.59	35.93	-13.93	76.25	25.48	27.53	36.81	31.27
15	145.51	322.34	59.70	61.07	172.50	124.89	129.80	99.23	3.84	80.19	75.22	51.31
16	-73.54	59.70	185.06	95.64	6.23	21.67	81.70	29.55	-58.77	43.30	54.39	77.60
17	43.01	61.07	95.64	183.62	76.33	92.39	28.73	85.72	44.57	63.86	91.45	98.95
18	135.59	172.50	6.23	76.33	155.14	100.73	39.91	68.71	10.95	29.86	50.48	44.90
19	35.93	124.89	21.67	92.39	100.73	183.21	82.62	89.28	53.86	27.32	108.59	65.23
20	-13.93	129.80	81.70	28.73	39.91	82.62	171.80	18.25	2.41	18.87	7.81	8.08
21	76.25	99.23	29.55	85.72	68.71	89.28	18.25	189.97	-5.67	-9.26	69.80	59.29
22	25.48	3.84	-58.77	44.57	10.95	53.86	2.41	-5.67	129.13	64.96	51.60	-22.05
23	27.53	80.19	43.30	63.86	29.86	27.32	18.87	-9.26	64.96	123.73	111.51	37.73
24	36.81	75.22	54.39	91.45	50.48	108.59	7.81	69.80	51.60	111.51	235.41	84.52
25	31.27	51.31	77.60	98.95	44.90	65.23	8.08	59.29	-22.05	37.73	84.52	127.91

Table 23 (cont'd)

σ_{ij}	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**
1	-30.78	83.66	88.78	131.05	75.71	87.03	73.71	90.56	57.04	56.39	55.67	51.12	55.98
2	149.30	42.53	168.93	182.02	90.97	84.43	173.62	100.65	131.76	119.26	57.91	68.36	97.75
3	15.03	41.29	43.72	91.00	53.27	65.72	45.97	48.94	64.50	36.15	38.25	47.66	52.24
4	106.58	-79.67	28.12	108.13	114.01	67.76	96.08	26.35	120.64	40.39	14.27	3.04	80.39
5	-25.86	73.96	24.79	-17.23	-60.30	-18.66	1.36	13.76	-0.90	-38.08	15.60	53.42	-5.63
6	104.75	13.73	115.35	156.60	101.55	93.91	172.79	29.51	99.88	104.84	23.12	28.71	133.08
7	123.46	-4.52	28.34	85.36	51.59	4.77	114.30	-26.74	44.97	91.82	19.07	60.91	120.72
8	74.72	43.58	108.99	107.65	51.14	59.88	128.03	25.68	89.55	44.36	7.37	43.96	95.97
9	58.81	60.01	127.20	145.53	76.92	66.20	112.12	83.44	74.22	61.04	78.76	97.53	98.20
10	-2.05	80.62	62.56	135.61	84.71	75.29	90.75	63.87	97.59	54.30	17.21	51.95	74.15
11	-3.70	52.27	26.56	125.94	82.05	101.78	54.32	73.92	88.29	47.08	64.19	47.97	31.13
12	42.87	35.20	22.14	166.28	119.33	88.66	81.70	92.02	117.15	86.24	78.49	77.75	67.37
13	6.68	109.41	23.36	143.05	35.56	107.43	72.79	130.21	89.76	97.87	133.63	83.81	-20.99
14	144.11	-41.26	50.44	153.48	145.37	67.52	107.37	10.32	78.29	116.57	44.22	44.68	109.51
15	182.45	41.68	180.02	220.07	104.05	110.52	207.40	114.70	134.26	130.79	138.96	124.00	118.59
16	-58.57	124.57	83.13	26.89	-51.72	57.78	3.79	90.85	31.30	27.55	2.05	-6.53	-50.49
17	-30.44	98.31	76.42	131.69	66.26	112.87	57.25	112.70	97.49	45.70	49.44	31.55	-4.65
18	59.74	76.72	83.75	157.04	76.89	86.43	111.03	77.05	91.50	70.78	97.30	108.18	77.60
19	70.50	51.89	98.94	132.71	64.40	42.01	123.41	80.18	98.67	24.20	83.29	111.19	71.27
20	43.74	41.46	161.55	42.67	-8.08	31.66	53.40	65.17	49.08	-9.11	2.38	28.46	42.40
21	61.65	-8.57	66.52	75.54	65.83	-12.73	99.96	25.42	-11.47	38.24	68.53	36.15	-3.96
22	32.90	-51.42	22.08	96.22	111.10	64.40	62.80	28.11	78.50	9.89	29.02	18.38	77.34
23	37.72	2.99	51.75	142.70	107.08	115.75	63.99	95.59	114.41	82.86	48.60	20.67	65.15
24	47.68	21.84	41.53	175.20	137.73	77.67	108.97	80.67	108.78	108.75	51.78	61.52	108.13
25	23.46	51.45	35.17	64.49	16.57	42.94	3.08	95.84	69.83	46.49	35.94	23.49	-46.36

Table 23 (cont'd)

σ_{ij}	**39**	**40**	**41**	**42**	**43**	**44**	**45**	**46**	**47**	**48**	**49**	**50**
1	151.65	207.94	79.42	49.31	50.24	76.03	66.85	74.55	-5.67	33.38	125.02	54.34
2	243.72	170.43	175.91	-41.38	120.00	98.62	96.27	69.60	-18.84	73.31	169.54	37.17
3	63.86	93.17	34.91	-7.66	0.60	17.41	25.57	4.97	35.97	27.64	68.19	54.35
4	131.83	35.62	90.50	-23.07	0.92	-76.46	28.79	-18.22	68.68	70.50	18.45	84.61
5	27.56	18.76	-34.10	-15.80	21.79	33.81	19.01	12.79	7.40	-13.37	39.98	-24.60
6	176.39	96.11	164.70	-27.10	38.89	65.01	69.52	41.72	-25.62	66.82	108.25	51.31
7	87.45	74.44	106.14	-33.11	3.13	45.69	-24.31	-13.34	-1.28	25.62	136.00	-17.13
8	164.41	75.45	96.71	-69.99	19.23	30.45	68.57	26.53	16.02	36.02	104.93	54.74
9	165.31	165.94	97.65	45.47	39.55	126.00	73.78	46.29	8.76	49.25	155.76	14.44
10	211.95	242.26	83.61	-52.34	30.62	-31.49	40.54	51.00	57.34	19.00	169.15	109.85
11	75.76	128.84	54.33	-14.69	2.30	-31.84	27.54	10.89	67.04	23.88	49.48	122.49
12	162.39	241.76	94.84	2.79	40.73	-33.55	6.65	20.96	80.63	43.16	144.45	105.87
13	71.28	172.91	94.00	14.78	131.41	29.95	25.39	57.77	2.31	32.45	55.11	80.37
14	63.34	81.64	127.74	-45.27	-64.81	4.51	-7.98	-39.55	65.04	24.13	102.52	91.99
15	187.76	115.75	197.91	23.18	97.90	169.46	104.58	46.59	-12.89	89.95	135.25	26.74
16	57.74	84.52	8.48	-39.03	136.75	65.30	71.74	89.17	-79.95	14.58	29.25	5.35
17	140.52	180.24	64.90	-46.00	59.18	-23.87	72.14	62.16	42.82	17.18	70.50	140.67
18	118.57	147.23	95.44	-9.15	14.68	58.47	44.38	16.39	52.88	29.05	124.38	78.00
19	220.78	168.54	95.63	5.00	47.92	37.16	69.13	39.24	67.40	35.31	146.73	52.74
20	110.89	7.58	20.37	1.50	61.21	132.99	113.78	47.37	-42.43	51.11	56.97	-43.89
21	133.51	88.12	133.02	38.61	16.86	45.29	45.38	54.94	7.91	-26.47	62.43	40.43
22	103.33	63.21	53.23	41.12	-39.70	-38.87	30.80	-10.95	71.90	52.04	24.69	75.31
23	106.26	140.46	78.14	31.34	62.22	14.38	40.45	27.49	11.91	83.24	61.32	63.69
24	228.94	297.29	130.91	31.82	65.26	1.47	13.67	52.69	36.98	59.21	200.16	67.78
25	72.22	122.51	26.76	-54.52	94.95	-24.82	20.90	33.65	23.12	10.95	54.69	57.74

Table 23 (cont'd)

σ_{ij}	1	2	3	4	5	6	7	8	9	10	11	12	13
26	-30.78	149.30	15.03	106.58	-25.86	104.75	123.46	74.72	58.81	-2.05	-3.70	42.87	6.68
27	83.66	42.53	41.29	-79.67	73.96	13.73	-4.52	43.58	60.01	80.62	52.27	35.20	109.41
28	88.78	168.93	43.72	28.12	24.79	115.35	28.34	108.99	127.20	62.56	26.56	22.14	23.36
29	131.05	182.02	91.00	108.13	-17.23	156.60	85.36	107.65	145.53	135.61	125.94	166.28	143.05
30	75.71	90.97	53.27	114.01	-60.30	101.55	51.59	51.14	76.92	84.71	82.05	119.33	35.56
31	87.03	84.43	65.72	67.76	-18.66	93.91	4.77	59.88	66.20	75.29	101.78	88.66	107.43
32	73.71	173.62	45.97	96.08	1.36	172.79	114.30	128.03	112.12	90.75	54.32	81.70	72.79
33	90.56	100.65	48.94	26.35	13.76	29.51	-26.74	25.68	83.44	63.87	73.92	92.02	130.21
34	57.04	131.76	64.50	120.64	-0.90	99.88	44.97	89.55	74.22	97.59	88.29	117.15	89.76
35	56.39	119.26	36.15	40.39	-38.08	104.84	91.82	44.36	61.04	54.30	47.08	86.24	97.87
36	55.67	57.91	38.25	14.27	15.60	23.12	19.07	7.37	78.76	17.21	64.19	78.49	133.63
37	51.12	68.36	47.66	3.04	53.42	28.71	60.91	43.96	97.53	51.95	47.97	77.75	83.81
38	55.98	97.75	52.24	80.39	-5.63	133.08	120.72	95.97	98.20	74.15	31.13	67.37	-20.99
39	151.65	243.72	63.86	131.83	27.56	176.39	87.45	164.41	165.31	211.95	75.76	162.39	71.28
40	207.94	170.43	93.17	35.62	18.76	96.11	74.44	75.45	165.94	242.26	128.84	241.76	172.91
41	79.42	175.91	34.91	90.50	-34.10	164.70	106.14	96.71	97.65	83.61	54.33	94.84	94.00
42	49.31	-41.38	-7.66	-23.07	-15.80	-27.10	-33.11	-69.99	45.47	-52.34	-14.69	2.79	14.78
43	50.24	120.00	0.60	0.92	21.79	38.89	3.13	19.23	39.55	30.62	2.30	40.73	131.41
44	76.03	98.62	17.41	-76.46	33.81	65.01	45.69	30.45	126.00	-31.49	-31.84	-33.55	29.95
45	66.85	96.27	25.57	28.79	19.01	69.52	-24.31	68.57	73.78	40.54	27.54	6.65	25.39
46	74.55	69.60	4.97	-18.22	12.79	41.72	-13.34	26.53	46.29	51.00	10.89	20.96	57.77
47	-5.67	-18.84	35.97	68.68	7.40	-25.62	-1.28	16.02	8.76	57.34	67.04	80.63	2.31
48	33.38	73.31	27.64	70.50	-13.37	66.82	25.62	36.02	49.25	19.00	23.88	43.16	32.45
49	125.02	169.54	68.19	18.45	39.98	108.25	136.00	104.93	155.76	169.15	49.48	144.45	55.11
50	54.34	37.17	54.35	84.61	-24.60	51.31	-17.13	54.74	14.44	109.85	122.49	105.87	80.37

Table 23 (cont'd)

σ_{ij}	14	15	16	17	18	19	20	21	22	23	24	25
26	144.11	182.45	-58.57	-30.44	59.74	70.50	43.74	61.65	32.90	37.72	47.68	23.46
27	-41.26	41.68	124.57	98.31	76.72	51.89	41.46	-8.57	-51.42	2.99	21.84	51.45
28	50.44	180.02	83.13	76.42	83.75	98.94	161.55	66.52	22.08	51.75	41.53	35.17
29	153.48	220.07	26.89	131.69	157.04	132.71	42.67	75.54	96.22	142.70	175.20	64.49
30	145.37	104.05	-51.72	66.26	76.89	64.40	-8.08	65.83	111.10	107.08	137.73	16.57
31	67.52	110.52	57.78	112.87	86.43	42.01	31.66	-12.73	64.40	115.75	77.67	42.94
32	107.37	207.40	3.79	57.25	111.03	123.41	53.40	99.96	62.80	63.99	108.97	3.08
33	10.32	114.70	90.85	112.70	77.05	80.18	65.17	25.42	28.11	95.59	80.67	95.84
34	78.29	134.26	31.30	97.49	91.50	98.67	49.08	-11.47	78.50	114.41	108.78	69.83
35	116.57	130.79	27.55	45.70	70.78	24.20	-9.11	38.24	9.89	82.86	108.75	46.49
36	44.22	138.96	2.05	49.44	97.30	83.29	2.38	68.53	29.02	48.60	51.78	35.94
37	44.68	124.00	-6.53	31.55	108.18	111.19	28.46	36.15	18.38	20.67	61.52	23.49
38	109.51	118.59	-50.49	-4.65	77.60	71.27	42.40	-3.96	77.34	65.15	108.13	-46.36
39	63.34	187.76	57.74	140.52	118.57	220.78	110.89	133.51	103.33	106.26	228.94	72.22
40	81.64	115.75	84.52	180.24	147.23	168.54	7.58	88.12	63.21	140.46	297.29	122.51
41	127.74	197.91	8.48	64.90	95.44	95.63	20.37	133.02	53.23	78.14	130.91	26.76
42	-45.27	23.18	-39.03	-46.00	-9.15	5.00	1.50	38.61	41.12	31.34	31.82	-54.52
43	-64.81	97.90	136.75	59.18	14.68	47.92	61.21	16.86	-39.70	62.22	65.26	94.95
44	4.51	169.46	65.30	-23.87	58.47	37.16	132.99	45.29	-38.87	14.38	1.47	-24.82
45	-7.98	104.58	71.74	72.14	44.38	69.13	113.78	45.38	30.80	40.45	13.67	20.90
46	-39.55	46.59	89.17	62.16	16.39	39.24	47.37	54.94	-10.95	27.49	52.69	33.65
47	65.04	-12.89	-79.95	42.82	52.88	67.40	-42.43	7.91	71.90	11.91	36.98	23.12
48	24.13	89.95	14.58	17.18	29.05	35.31	51.11	-26.47	52.04	83.24	59.21	10.95
49	102.52	135.25	29.25	70.50	124.38	146.73	56.97	62.43	24.69	61.32	200.16	54.69
50	91.99	26.74	5.35	140.67	78.00	52.74	-43.89	40.43	75.31	63.69	67.78	57.74

Table 23 (cont'd)

σ_{ij}	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**
26	276.94	-126.60	65.05	119.38	105.42	5.22	176.62	-14.94	88.57	119.18	54.19	44.71	115.09
27	-126.60	209.89	37.97	42.34	-68.01	75.15	-2.34	79.86	40.54	-0.11	30.83	65.91	-19.26
28	65.05	37.97	186.50	112.43	54.04	74.97	85.98	104.57	80.95	28.80	41.19	42.21	53.75
29	119.38	42.34	112.43	288.92	196.58	174.22	203.57	125.74	188.89	155.24	128.10	115.41	163.88
30	105.42	-68.01	54.04	196.58	203.16	99.42	128.72	51.16	109.19	106.58	63.29	41.19	137.95
31	5.22	75.15	74.97	174.22	99.42	163.13	73.75	117.95	141.02	84.57	62.01	45.46	69.74
32	176.62	-2.34	85.98	203.57	128.72	73.75	258.01	9.45	106.86	132.45	101.76	102.92	169.44
33	-14.94	79.86	104.57	125.74	51.16	117.95	9.45	178.00	116.34	28.05	87.22	55.69	-21.63
34	88.57	40.54	80.95	188.89	109.19	141.02	106.86	116.34	177.60	96.15	69.35	72.36	95.96
35	119.18	-0.11	28.80	155.24	106.58	84.57	132.45	28.05	96.15	147.63	46.84	37.09	103.05
36	54.19	30.83	41.19	128.10	63.29	62.01	101.76	87.22	69.35	46.84	138.45	99.82	22.33
37	44.71	65.91	42.21	115.41	41.19	45.46	102.92	55.69	72.36	37.09	99.82	119.74	71.60
38	115.09	-19.26	53.75	163.88	137.95	69.74	169.44	-21.63	95.96	103.05	22.33	71.60	217.90
39	109.59	40.19	151.15	246.12	167.61	101.70	218.88	107.81	153.20	97.83	101.28	123.01	164.91
40	-37.96	143.75	69.62	271.04	176.85	161.00	117.55	160.92	164.30	123.99	102.79	132.66	137.42
41	189.03	-39.32	71.35	204.53	153.41	68.12	244.69	13.61	93.37	154.59	101.95	74.10	142.56
42	-27.58	-59.64	18.60	18.95	56.27	-14.36	4.24	33.80	-39.68	-37.59	70.92	25.19	6.47
43	25.56	72.40	63.46	51.77	-33.13	50.87	28.97	111.32	69.93	43.95	56.21	30.13	-48.41
44	38.44	30.98	137.30	45.54	0.17	14.34	65.83	50.35	-1.90	11.55	47.38	46.39	43.86
45	3.97	39.56	131.49	66.58	24.02	60.42	41.42	94.07	50.96	-9.96	41.06	24.63	3.21
46	-35.98	59.09	58.56	37.37	0.62	29.08	26.26	57.77	10.07	7.44	28.77	15.35	-15.29
47	9.18	-6.73	-21.20	66.50	70.00	31.67	12.81	28.49	58.84	-3.09	39.66	50.46	32.56
48	65.29	-12.49	63.79	101.40	70.82	76.56	59.71	66.41	94.42	51.81	35.37	28.20	69.45
49	60.82	77.55	78.91	187.78	119.78	72.06	143.23	50.35	109.55	106.43	43.91	114.40	176.79
50	-11.82	53.25	7.80	139.67	101.23	116.60	59.01	67.84	100.94	60.71	51.45	33.59	32.83

Table 23 (cont'd)

σ_{ij}	39	40	41	42	43	44	45	46	47	48	49	50
26	109.59	-37.96	189.03	-27.58	25.56	38.44	3.97	-35.98	9.18	65.29	60.82	-11.82
27	40.19	143.75	-39.32	-59.64	72.40	30.98	39.56	59.09	-6.73	-12.49	77.55	53.25
28	151.15	69.62	71.35	18.60	63.46	137.30	131.49	58.56	-21.20	63.79	78.91	7.80
29	246.12	271.04	204.53	18.95	51.77	45.54	66.58	37.37	66.50	101.40	187.78	139.67
30	167.61	176.85	153.41	56.27	-33.13	0.17	24.02	0.62	70.00	70.82	119.78	101.23
31	101.70	161.00	68.12	-14.36	50.87	14.34	60.42	29.08	31.67	76.56	72.06	116.60
32	218.88	117.55	244.69	4.24	28.97	65.83	41.42	26.26	12.81	59.71	143.23	59.01
33	107.81	160.92	13.61	33.80	111.32	50.35	94.07	57.77	28.49	66.41	50.35	67.84
34	153.20	164.30	93.37	-39.68	69.93	-1.90	50.96	10.07	58.84	94.42	109.55	100.94
35	97.83	123.99	154.59	-37.59	43.95	11.55	-9.96	7.44	-3.09	51.81	106.43	60.71
36	101.28	102.79	101.95	70.92	56.21	47.38	41.06	28.77	39.66	35.37	43.91	51.45
37	123.01	132.66	74.10	25.19	30.13	46.39	24.63	15.35	50.46	28.20	114.40	33.59
38	164.91	137.42	142.56	6.47	-48.41	43.86	3.21	-15.29	32.56	69.45	176.79	32.83
39	358.14	302.93	204.90	43.66	66.35	71.56	100.14	77.35	57.84	82.46	243.54	86.99
40	302.93	491.87	123.67	36.82	65.77	8.87	38.50	82.32	93.73	63.30	306.40	154.85
41	204.90	123.67	264.38	18.54	36.42	51.36	28.79	30.97	3.66	54.09	124.87	66.20
42	43.66	36.82	18.54	200.83	-14.02	85.00	37.68	30.34	-0.42	22.60	-11.40	-45.27
43	66.35	65.77	36.42	-14.02	167.74	52.12	55.09	62.87	-48.85	44.14	16.35	-7.73
44	71.56	8.87	51.36	85.00	52.12	191.11	92.66	54.77	-73.51	39.06	51.10	-79.61
45	100.14	38.50	28.79	37.68	55.09	92.66	112.97	55.57	-13.47	42.60	16.28	15.49
46	77.35	82.32	30.97	30.34	62.87	54.77	55.57	63.55	-30.66	7.54	34.50	9.04
47	57.84	93.73	3.66	-0.42	-48.85	-73.51	-13.47	-30.66	110.45	7.31	50.06	83.40
48	82.46	63.30	54.09	22.60	44.14	39.06	42.60	7.54	7.31	77.14	46.94	17.39
49	243.54	306.40	124.87	-11.40	16.35	51.10	16.28	34.50	50.06	46.94	279.55	53.89
50	86.99	154.85	66.20	-45.27	-7.73	-79.61	15.49	9.04	83.40	17.39	53.89	164.18

Table 24. List of stocks used in Section 5.2.3

ACIBADEM SAĞLIK	BANVİT	DOĞAN HOLDİNG
ADANA ÇİMENTO (A)	BERDAN TEKSTİL	DOĞAN YAYIN HOLDİNG
ADANA ÇİMENTO (B)	BISAŞ TEKSTİL	DYO
ADANA ÇİMENTO (C)	BOLU ÇİMENTO	ECZACIBAŞI İLAÇ
ADEL KALEMCİLİK	BOSSA	ECZACIBAŞI YAPI
ANADOLU EFES	BOYNER	EGE ENDÜSTRİ
AFYON ÇİMENTO	BRİSA	EGE GÜBRE
AKAL TEKSTİL	BİRLİK MENSUCAT	EGE SERAMİK
AKÇANSA	BOROVA YAPI	EMEK ELEKTRİK
AK ENERJİ	BSH EV ALETLERI	EMİNİŞ AMBALAJ
AKSA	BATISÖKE ÇİMENTO	ENKA İNŞAAT
AKSU ENERJİ	BATI ÇİMENTO	EGEPLAST
ALARKO HOLDİNG	BURSA ÇİMENTO	ERBOSAN
ALARKO CARRIER	ÇBS BOYA	EREĞLİ DEMİR ÇELİK
ALCATEL LUCENT TELETAŞ	ÇELİK HAT	ERSU GIDA
ALKİM KAĞIT	ÇEMTAŞ	ESCORT TEKNOLOJİ
ALKİM KİMYA	CEYLAN GİYİM	ESEM
ALTINYILDIZ	ÇİMSA	FENİŞ ALÜMİNYUM
ALTINYAĞ	ÇELEBİ	F-M İZMİT PİSTON
ANADOLU CAM	ÇİMBETON	FRIGO
ARÇELIK	ÇİMENTAŞ	FORD OTOSAN
ARENA BİLGİSAYAR	COMPONENTO	FAVORİ DİNLENME YERLERİ
ARSAN TEKSTİL	DARDANEL	GEDİZ İPLİK
ASELSAN	DENİZLİ CAM	GENTAŞ
LAFARGE ASLAN ÇİMENTO	DENTAŞ AMBALAJ	GOLDAŞ
ANADOLU ISUZU	DERİMOD	GÖLTAŞ ÇİMENTO
AKIN TEKSTİL	DEVA HOLDİNG	GOODYEAR
ALTINYUNUS ÇEŞME	DOĞAN GAZETE	GSD HOLDING
AYEN ENERJİ	DİTAŞ DOĞAN	GÜBRE FABRIKALARI
AYGAZ	DEMİSAŞ DÖKÜM	HEKTAŞ
BAGFAŞ	DOĞAN BURDA	HÜRRİYET GAZETECİLİK
BAK AMBALAJ	DOĞUSAN	

Table 25. List of stocks used in Section 5.4

ACIBADEM SAĞLIK	BOSCH FREN	DOĞAN HOLDİNG
ADANA ÇİMENTO (A)	BISAŞ TEKSTİL	DURAN DOĞAN BASIM
ADANA ÇİMENTO (B)	BOLU ÇİMENTO	DOĞAN YAYIN HOLDİNG
ADANA ÇİMENTO (C)	BOSSA	DYO
ADEL KALEMCİLİK	BOYNER	ECZACIBAŞI İLAÇ
ANADOLU EFES	BRİSA	ECZACIBAŞI YAPI
AFYON ÇİMENTO	BİRLİK MENSUCAT	EGE ENDÜSTRİ
AKAL TEKSTİL	BOROVA YAPI	EGE GÜBRE
AKÇANSA	BORUSAN	EGE PROFİL
AK ENERJİ	BSH EV ALETLERI	EGE SERAMİK
AKSA	BATISÖKE ÇİMENTO	EMEK ELEKTRİK
AKSU ENERJİ	BATI ÇİMENTO	EMİNİŞ AMBALAJ
ALARKO HOLDİNG	BURSA ÇİMENTO	ENKA İNŞAAT
ALARKO CARRIER	BURÇELİK	EGEPLAST
ALCATEL LUCENT TELETAŞ	ÇBS BOYA	ERBOSAN
ALKİM KAĞIT	ÇELİK HAT	EREĞLİ DEMİR ÇELİK
ALKİM KİMYA	ÇEMTAŞ	ERSU GIDA
ALTINYILDIZ	CEYLAN GİYİM	ESCORT TEKNOLOJİ
ALTINYAĞ	ÇİMSA	ESEM
ANADOLU CAM	ÇELEBİ	FENİŞ ALÜMİNYUM
ARÇELIK	ÇİMBETON	F-M İZMİT PİSTON
ARENA BİLGİSAYAR	ÇİMENTAŞ	FRIGO
ARSAN TEKSTİL	COMPONENTO	FORD OTOSAN
ASELSAN	DARDANEL	FAVORİ DİNLENME YERLERİ
LAFARGE ASLAN ÇİMENTO	DENİZLİ CAM	GEDİZ İPLİK
ANADOLU ISUZU	DENTAŞ AMBALAJ	GENTAŞ
AKIN TEKSTİL	DERİMOD	GOLDAŞ
ALTINYUNUS ÇEŞME	DEVA HOLDİNG	GÖLTAŞ ÇİMENTO
AYEN ENERJİ	DOĞAN GAZETE	GOODYEAR
AYGAZ	DİTAŞ DOĞAN	GSD HOLDING
BAGFAŞ	DEMİSAŞ DÖKÜM	GÜBRE FABRIKALARI
BAK AMBALAJ	DOĞAN BURDA	HEKTAŞ
BANVİT	DOĞUSAN	HÜRRİYET GAZETECİLİK
BERDAN TEKSTİL		

APPENDIX B

COMPARISON OF VARIANCE, MAD AND MUA WITH STOCKS FROM ISE

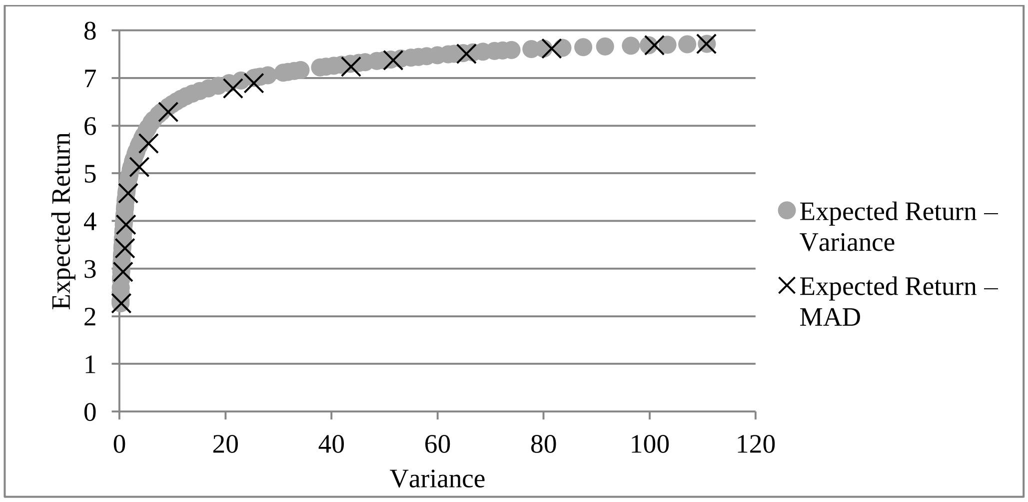


Figure 45. Comparing expected return–variance and expected return–MAD efficient portfolios in expected return and variance

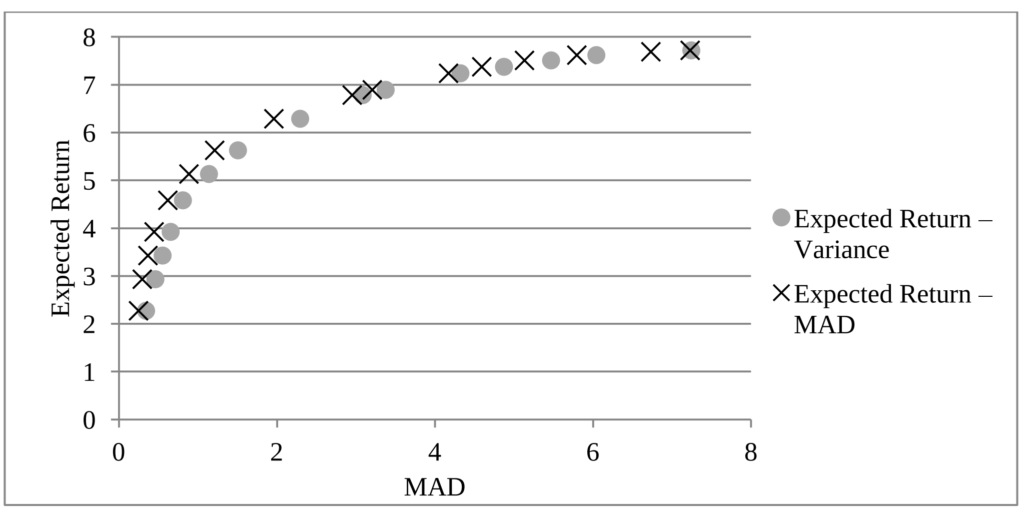


Figure 46. Comparing expected return–variance and expected return–MAD efficient portfolios in expected return and MAD

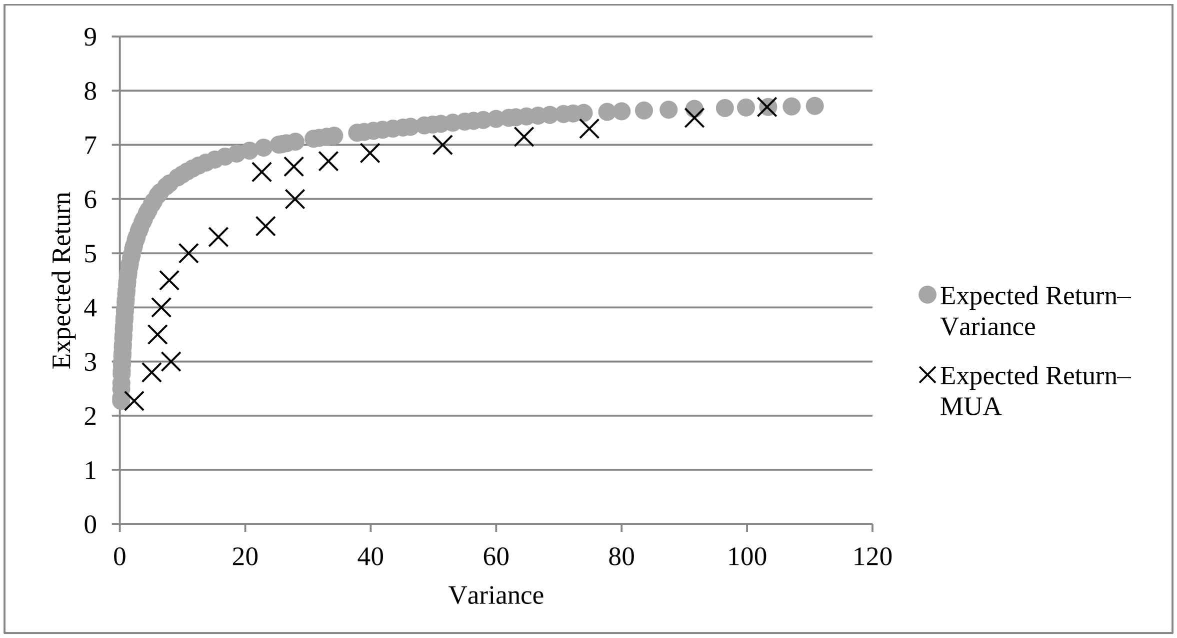


Figure 47. Comparing expected return–variance and expected return–MUA efficient portfolios in expected return and variance

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Tuncer Şakar, Ceren
Nationality: Turkish (TC)
Date and Place of Birth: 16 June 1981, Antalya
Marital Status: Married
Phone: +90 312 217 15 75
email: cerentuncer@yahoo.com

EDUCATION

Degree	**Institution**	**Year of Graduation**
M.S.	METU Industrial Engineering	2006
B.S.	METU Business Administration	2003
High School	Antalya Anatolian High School	1999

WORK EXPERIENCE

Year	**Place**	**Enrollment**
2009- Present	METU Industrial Engineering	Research Assistant
2006-2009	BİLGİ GIS Ltd.	Administrative and Financial Specialist
2005-2006	METU Industrial Engineering	Research Assistant
2004-2005	ASELSAN A.Ş.	Accounting Junior Specialist

PUBLICATIONS

Tuncer Şakar, C. and Köksalan M. (2013), "Effects of Multiple Criteria on Portfolio Optimization," forthcoming in *the International Journal of Information Technology & Decision Making*.

Tuncer Şakar, C. and Köksalan M. (2013), "A Stochastic Programming Approach to Multicriteria Portfolio Optimization," *Journal of Global Optimization*, 57(2), p. 299-314.

Köksalan, M. and Tuncer C. (2009), “A DEA-Based Approach to Ranking Multi-Criteria Alternatives,” *International Journal of Information Technology & Decision Making*, 8, p. 29-54.

CONFERENCES AND SEMINARS

Tuncer Şakar, C. and Köksalan M., “An Interactive Approach to Stochastic Programming-based Portfolio Optimization,” International IIE Conference, İstanbul, Turkey, June 2013.

Tuncer Şakar, C. and Köksalan M., “Effects of Multiple Criteria on Portfolio Optimization,” Department of Industrial Engineering, Middle East Technical University, Ankara, Turkey, October 2012.

Tuncer Şakar, C. and Köksalan M., “Effects of Multiple Criteria and Different Planning Horizons on Portfolio Optimization,”□21st International Symposium on Mathematical Programming (ISMP), Berlin, Germany, August 2012.

Tuncer Şakar, C. and Köksalan M., “A Stochastic Programming Approach to Multicriteria Portfolio Optimization,” INFORMS Annual Meeting, Charlotte, North Carolina, USA, November 2011.

Tuncer Şakar, C. and Köksalan M., “A Stochastic Programming Approach to Multicriteria Portfolio Optimization,” 21st International Conference on MCDM, Jyväskylä, Finland, June 2011.

Tuncer Şakar, C., “A Preference-Based Evolutionary Algorithm to Solve Portfolio Optimization with Cardinality and Weight Constraints,” Preconference of the 21st International Conference on MCDM, Helsinki, Finland, June 2011.

Tuncer Şakar, C., “A Stochastic Programming Approach to Multicriteria Portfolio Optimization,” Aalto University, School of Economics, Helsinki, Finland, March 2011.

Tuncer Şakar, C. and Köksalan M., “A Multicriteria Decision Making Approach to Portfolio Optimization,” 20th International Conference on MCDM, Chengdu, China, June 2009.

FOREIGN LANGUAGES

Advanced English, Beginner German

HOBBIES

Free and Scuba Diving, Movies