The use of risk measures and its applications in portfolio optimisation

By

Resham Sivnarain

(11230712)

Submitted in partial fulfilment of the requirements for the degree

Magister Scientiae

In the Department of Mathematics and Applied Mathematics

In the Faculty of Natural and Agriculture Sciences

University of Pretoria

Pretoria

August 2017

Declaration

I, the undersigned, declare that this dissertation which I hereby submit for the degree Magister Scientiae at the University of Pretoria, is my own work and has not previously been submitted by me for any degree at this or any other tertiary institute.

Signature: ____________________

Name: Resham Sivnarain

Date:

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Abstract

The use of risk measures and its applications in portfolio optimisation

Resham Sivnarain

Magister Scientiae

In the Department of Mathematics and Applied Mathematics

University of Pretoria

2017

Supervisor: Prof. E. Maré

In this dissertation, we study the application of risk measures to portfolio optimisation. A risk measure is a functional over the set of random portfolio returns mappings $X \to \rho(X) \in \mathbb{R}$. We present the various risk measures in this dissertation within an axiomatic framework. Although Value-at-Risk (VaR) has been widely used, the Conditional-Value-at-Risk (CVaR) has become the more popular risk measure since it is a coherent and convex risk measure. We solve a CVaR based optimisation model that is used for portfolio optimisation and hedging a target portfolio. Additionally, we solve a CVaR based optimisation model with cost considerations included in the objective function. Further, we include alternative risk measures such as distortion, spectral, drawdown and coherent-distortion risk measures (CDRM) and develop optimisation problems for each risk measure as either the objective function or as a constraint in a linear programming problem. Since the 2008 crisis era, it has become important to note the universal agreement that financial assets have fat tails and that financial and investment managers must be able to account for it in their risk management strategies. We present fat-tail analysis for CVaR optimisation problems and perfom emperical risk analysis on the FTSE/JSE ALSI index.

Dedication

ॐ गम गणपतये नमः

Om Gam Ganapataye Namaha

This thesis is dedicated to my father, Mr. P. Sivnarain

Acknowledgements

I would like to thank God for giving me this opportunity. I would like to convey my sincere gratitude to my supervisor, Prof. Eben Mare for his patience and words of advise during the course. Finally, I would like to thank all my loved ones for their support and kindness.

Table of Contents

Table of Figures

Table of Tables

List of Abbreviations

Abbreviation	Description
AD	Absolute drawdown risk measure
AvDD	Average drawdown risk measure
CDD	Conditional drawdown risk measure
CDRM	Coherent distortion risk measure
CVaR	Conditional Value-at-Risk
ES	Expected shortfall
EVT	Extreme value theory
MaxDD	Maximum drawdown risk measure
SSD	Second order stochastic dominance
VaR	Value-at-Risk

List of Symbols

Symbol	Description
ρ	Convex measure of risk.
α	Loss not exceeding a threshold of α.
β	Confidence level.
X	Given portfolio returns.
F_X	Corresponding distribution function for portfolio returns X.
M_ϕ	Estimator of a spectral measure.
$\psi(x,\zeta)$	Resulting loss distribution function for loss $z = f(x,y)$.
ϕ_i	Admissible sequence used in spectral risk measures.
S_0	Initial asset price.
δV	Change in value of derivative.
g	Distortion function.
l	Lower bound limit used in linear programming problem.
u	Upper bound limit used in linear programming problem.
c_i	Weight cost parameter.
$VaR_\alpha(X)$	Value at risk for a portfolio of returns X, where the loss does not exceed a given α.
$CVaR_\alpha(X)$	Conditional value at risk for a portfolio of returns X, where the loss does not exceed a given α.
$\left\|\overline{CVaR}\right\|$	Optimal CVaR with no cost consideration.

Nomenclature

Value-at-Risk

Is a statistical technique used to measure and quantify the level of financial risk for an investment portfolio over a specific time frame.

Spectral risk measure

A Spectral risk measure is a risk measure given as a weighted average of outcomes where bad outcomes are, typically, included with larger weights. Aspectral risk measure is a function of portfolio returns and outputs the amount of the numeraire (typically a currency) to be kept in reserve.

Ill-posedness

A well-posed mathematical problem should have the properties that:

i. A solution exists
ii. The solution is unique
iii. The solution's behavior changes continuously with the initial conditions.

Problems that are not well-posed are termed ill-posed.

Efficient frontier

The efficient frontier is the set of optimal portfolios that offers the highest expected return for a defined level of risk or the lowest risk for a given level of expected return.

Chapter 1. Introduction

This chapter serves as an introduction to risk measures and its applications to portfolio optimisation. The need for portfolio optimisation with various risk measures is introduced on a conceptual level with a detailed literature review also presented in this section.

1.1. Background and introduction

Optimal portfolio allocation is a longstanding problem in both practical portfolio management and academic research. Whether you are an investor, hedger, or fund manager among others striving to achieve an optimal portfolio is of paramount importance. The main objective in portfolio management is the trade-off between risk and return. Markowitz [24] was the first to study in-depth the problem of portfolio optimisation in the 1950's. He looked at maximizing portfolio expected return for a given level of risk or equivalently minimizing risk for given expected return. The classical work of Markowitz used variance as the benchmark for risk measurement. This method had a major shortfall since it penalized equally, regardless of the downside risk or upside potential. Over the decades, an evolution of risk measurement took place that has been proposed to be used in portfolio optimisation. Some of these risk measures include Value-at-Risk (VaR), partial moments, safety first principle, skewness and kurtosis, and Conditional Value-at-Risk (CVaR). Many others have also been developed and shall also be considered in this thesis.

Another shortfall of the Markowitz theory was that it applied well to linear financial instruments. In modern' time, mixed portfolios that comprise both linear financial instruments (Stocks, etc.) and nonlinear financial instruments (derivatives, etc.) have become the norm within portfolio management. This is due to the fact that derivative instruments are no longer considered as hedging instruments but now are considered as investment instruments. For example, options are the derivative instruments which can increase the liquidity and flexibility of return from the investment and at the same time it can be considered as assets to be invested.

In this research, we shall look at risk measures and how they influence the characteristics of an optimal portfolio that consists of linear and nonlinear financial instruments. In the next subsection, we shall give a very distinct view from the financial engineer's perspective on risk measures and portfolio optimisation.

1.1.1. A financial engineer's perspective

Before one can perform anything with an optimal portfolio, we need to understand or try to apply a framework for optimal portfolios. From a theoretical perspective, there are two well-known approaches to manage portfolio performance namely the Expected Utility Theory and Risk Management [11]. The financial engineer is very much interested in building a model to best address optimal portfolios. Let's suppose a given time interval $[0,T]$ is partitioned into N subintervals $[t_{k-1},t_k]$, $k=1,\dots,N$ by the set of points $\{t_0=0,t_1,t_2,\dots,t_N=T\}$. Let's also suppose there are m risky assets with rates of returns determined by a random vector $(r_1(t_k),r_2(t_k),\dots,r_m(t_k))$ for times t_k. We introduce a risk-free asset with constant rate of return $r_i(t_k)=\dfrac{p_i(t_k)}{p_i(t_{k-1})}-1.$

Consequently, a portfolio formed of the m risky assets and the risk-free instrument is determined by the vector of weights $x(t_k)=(x_0(t_k),x_1(t_k),\dots,x_m(t_k))$. The components of $x(t_k)$ satisfy the budget constraint,

$$\sum_{i=0}^{m} x_i(t_k)=1. \qquad 1.1$$

By definition, the rate of return of the portfolio at time moment t_k is

$$r_k^p(x(t_k))=r(t_k)\cdot x(t_k)=\sum_{i=0}^{m} r_i(t_k)x_i(t_k). \qquad 1.2$$

1.1.2 A framework for portfolio optimisation

In this subsection, we shall take a high level perspective of a framework for portfolio optimisation. Figure 1-1 below shows two very distinct but separate ways a portfolio manager can use optimisation for portfolio management. We shall consider two segments namely the, Expected Utility Theory aspect and General Risk Management framework. Each segment considers some optimisation problem, key characteristics and different risk measures that may be applied. This is a suggestion based on the work of Chekhlov et al. [11].

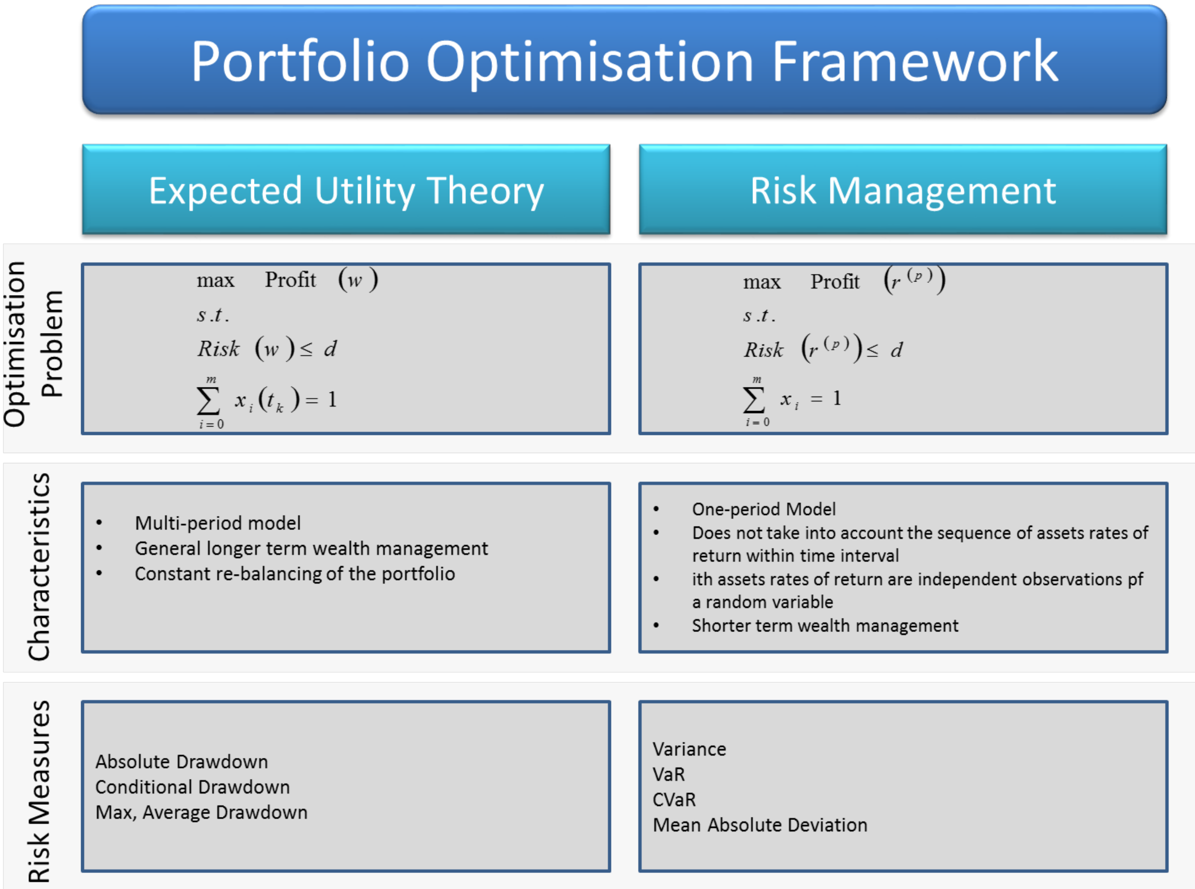


Figure 1-1: Portfolio optimisation framework

In this thesis, we shall primarily focus on the risk management segment where we use various different risk measures and apply it to the optimisation problem on the right of Figure 1-1. This means that constant rebalancing of portfolio weights will be more instantaneous (shorter time intervals) than longer time spans. Although drawdown measures are particularly applied in Expected Utility Theory, we shall apply drawdown measures in General Risk Management segment. This is to test the applicability of drawdown in a wider portfolio optimisation problem setting. The literature review is shown in the succeeding subsection.

1.2. Literature review

Derivative contracts are widely used by financial institutes and investors to achieve higher returns and to decrease a portfolio's funding costs. In addition, derivatives use has fundamentally changed the landscape for financial risk management. Derivatives that are used for either investment or risk management need a risk measure to be chosen to evaluate the performance of a portfolio. Over the past decades, various risk measures have been developed each providing its own benefits for portfolio performance measurement. Value-at-Risk (VaR) has grown to be the most popular standard benchmark [6] for firm wide measures of risk. For a given time horizon t and confidence level β, the Value-at-Risk of a given portfolio is the loss in the portfolios market value over the time horizon t that is exceeded with probability $1-\alpha$. However, as a risk measure, VaR has proven limitations. According to Artzner et al. [8], VaR lacks sub-additively and convexity.

Artzner et al. [8, 9] have defined the measures of risk and presented a unified framework for analysis, construction and implementation of risk measures. Artzner et al. [8] define concisely a coherent risk measure as one that satisfies the axioms of translation invariance, sub-additively, positive homogeneity, and monotonicity. Follmer et al. [16, 17, 18] extended the notion of coherent risk measures to convex risk measures. Follmer et al. [17] prove the corresponding extension of the representation theorem in terms of probability measures on the underlying space of scenarios and the representation theorem is closely related to the super-hedging duality under convex constraints [18, 20]. Wang et al. wanted to extend the notion of coherent risk measures and thus added two more axioms, which include law-invariance and comonotonic additivity, but lacked sub-addivity. Acerbi [1] studied the space of coherent risk measures obtained as certain expansions of coherent elementary basis measures. This new class of risk measures that relied on a spectrum was aptly called spectral risk measures. These spectral risk measures in addition to the axoims of Artzner et al. [8] include law-invariance and comonotonic additivity.

1.1.1. Conditional Value-at-Risk

An alternative measure of risk to the VaR is conditional Value-at-Risk (CVaR), which is also known as mean shortfall, expected shortfall and tail VaR [5, 6]. With a continuous distribution, for given time horizon t and confidence level β, CVaR is the conditional expectation of the loss above VaR for the time horizon t and the confidence level α. Based on this CVaR gives further information on the magnitude of the excess loss. CVaR has proven to have better properties than VaR [5, 6, 35, 36]. Pflug [29] has shown that CVaR is a coherent risk measure due to the properties mentioned above.

Rockafellar and Uryasev [35, 36] proposed a convex optimisation problem, which aims to compute an optimal minimum CVaR for a portfolio of derivative. Rockafella and Uryasev [35, 36] show that VaR is difficult to optimise when it is calculated from scenarios. Mauser and Rosen [26] and Mckay and Keefer [27] showed that VaR can be ill-behaved as a function of portfolio positions and can exhibit multiple local extrema to determine an optimal mix of positions. Uryasev [39] shows a simple description approach to minimizing the CVaR and optimisation problems with CVaR constraints. Outside the realm of finance, CVaR or similar measure, such as conditional expectation constraints and integrated chance constraints have been used before in stochastic programming literature [31]. Many numerical algorithms have been developed for solving stochastic optimisation problems [10, 14, 21, 22, 30, 31]. The main advantage of these algorithms is that they are able to make use of special mathematical features in the portfolio and can be readily combined with analytic and simulation based methods.

The most recent literature that captures further important work on derivative portfolio optimisation is Alexander et al. [5, 6]. Alexander et al. [6] look at the well-posedness of the CVaR and VaR optimisation selection problem where the investment universe includes derivatives. They illustrate that the CVaR/VaR optimisation problem for derivative portfolios typically has an infinite number of solutions if the derivative values are computed using delta-gamma approximations. They further investigate and illustrate that when derivative values are computed using more accurate methods such as analytic formulae, numerical partial differential equations, or Monte Carlo methods the CVaR/VaR optimisation problem for derivative portfolios remains ill-posed. The ill-posedness of the optimisation problem is the sense that there are many portfolios that have similar CVaR/VaR values to that of the optimal portfolio and any slight perturbations of the data can lead to significantly different optimal portfolios.

Alexander et al. [6] propose a CVaR optimisation problem with a convex programming problem that deals with modelling portfolio costs. The proportional cost model demonstrates that CVaR optimisation formulation with cost is limited to both transactional and management costs. They also demonstrate that CVaR investment portfolios using a suitably weighted cost parameter has smaller total trading positions, fewer instruments and comparable CVaR.

The typical method for solving CVaR optimisation problems make use of linear programming (LP) and most often a stochastic linear programming approach [6]. The linear programming problem is formed using Monte Carlo simulations and piecewise linear functions to approximate the continuous differentiable CVaR function. Alexander et al. [6] introduced a computation method based on a smoothing technique to efficiently solve the simulation based CVaR optimisation problem.

1.1.1. Tail risk and portfolio optimisation

Since the 2008 crisis era, it was important to note the universal agreement that financial assets have fat tails and that financial and investment managers must be able to account for it in their risk management strategies. In today's real world examples, it is necessary to assume a distributional hypothesis capable of describing both fat tails and asymmetry. From a practical and theoretical perspective, there have been several classes of distributions that have been used to capture fat tails in modelling. Rachev et al. [33, 34, 37, 38] conclude that the most popular is the Student's t distribution. Rachev et al. [33] also mention others that also form particular importance, namely extreme value distributions, stable distributions, operator stable distortions, the class of tempered stable distributions, and the class of infinitely divisible distributions. All these classes of models share one feature; that is they include normal distributions as a special or limiting case. The exception to this is the extreme value theory.

Stoyanov et al. [37, 38] considered the sensitivity of CVaR with respect to tail indexes such as Student t distribution's, degrees of freedom parameter. The sensitivity of CVaR findings from Stoyanov et al. [37, 38] are presented in this thesis without analytic proof. However, we do test the applicability of the findings to JSE data in Case Study 4 using simulation based methods and the draw the conclusion that the findings are indeed correct.

1.1.1. Alternative risk measures

Acerbi and Simonetti [2] studied spectral measures of risk in portfolio optimisation. They show that the minimization problem of a spectral measure is shown to be equivalent to the minimization of a suitable function that contains some additional analytic properties. Their study revealed that results of the classical risk-reward problem coincided with results of the unconstraint optimisation problem where a single suitable spectrum function was used. Adam et al. [3] study the use of spectral risk measures in portfolio optimisation under risk constraints and develop a comparative analysis of efficient portfolios.

Distortion risk measure [32] finds its origins in Yaari's 1987 paper on Dual theory of choice risk. Yaari's idea was to apply a distortion function on a distribution function. The definition of distortion risk measures makes reference on the theory to the Choquet integral. In 2000, the distortion risk measure [32] was applied to the insurance industry to solve a wide range of problems such as insurance premiums, capital requirements and capital allocation. Wang (see [32] for further details) has applied the distortion risk measure to price catastrophe bonds and Fabozzi and Tunaru (see [32] for further details) to price real estate derivatives. Rachev et al. [32] propose a new distortion risk measure, adding the asymmetric property to the existing properties. Rachev et al. further [32] extend the theory of the Choquet integral construction by using quadratic and power functions. Feng and Tan [15] introduce the theme of Coherent Distortion risk measures (CDRM) that are used in portfolio selection. CDRM comprise many risk measures such as CVaR, the Wang Transform measure and the proportional hazard measure [15].

A typical scenario where an investor is caught in a liquidity trap, such that he/she is unable to secure funding after an abrupt market decline (such as the 2008 financial crisis) is where maximum drawdown risk measure can be very useful. The analytic study of drawdown risk measure magnitudes has been studied in the applied probability theory. Taylor [19] showed the mathematical analysis of the maximum drawdown of the Brownian motion and was generalized by Lehoczky [19]. In 2012, Mijatovic and Pistorius analysed the drawdowns of spectrally negative Levy processes. In 2015, Landriault et al. [19] extended drawdown magnitude by studying the frequency rate of drawdown for the Brownian motion. Over a period of time, concepts such as "drawup" and reduction of drawdown in active portfolio management received much attention in mathematical finance research [19]. Chekhlov et al. [11] introduced a new one-parameter family of risk measures called Conditional Drawdown (CDD) or Conditional Drawdown at Risk (CDaR). In 2015, Goldberg and Mahmoud [19] formalized Conditional Expected Drawdown (CED), which can be understood as a tail mean of maximum drawdown distributions. Like CED, CDaR is a deviation measure however, unlike CED, CDaR focuses on all drawdowns rather than maximum drawdowns [11, 19].

1.3. Research objectives

Over the many decades of risk management, various risk measures have been developed. In 1952, Markowitz [24] introduced variance as a measure of risk. Since then many other risk measures have been developed that seemed to be mathematically superior to the simple variance risk measure. In this thesis, we shall consider the following as objectives for risk measures and its applications in portfolio optimisation:

(Obj. I). We shall present the theorems, properties and other propositions for theValue-at-Risk (VaR), Conditional Value-at-Risk (CVaR), spectral risk measures, distortion risk measures and Coherent Distortion risk measures from various literature resources. We shall use the Artzner et al. [8] axiomatic framework to form the basis of each risk measure's definition.

(Obj. II). We shall present for each of the risk measures above, an optimisation model(s) that can be solved to give the risk manager an optimal portfolio. The model solution will be based on risk or return or both as an objective function and constraint. These optimal problems shall be formulated for derivative based assets.

a. We wish to investigate the effects of solving a portfolio optimisation problem with each of the three risk measures, namely, CVaR, spectral risk measures and Coherent Distortion risk measures (CDRM) risk measure. These three risk measure have shown to have some popularity in both literature and in practice thus doing a comparative analysis among them may prove insightful.

(Obj. III). We shall also present the use of the CVaR risk measure to formulate an optimal hedging problem with derivative based contracts. This problem is typically constructed with a given target portfolio and a given hedging portfolio. The aim is to hedge the target portfolio with the given hedging portfolio. This problem also lends itself well to the practical setting of a risk manager who wishes to hedge risk of his current portfolio(s).

(Obj. IV). We shall present a CVaR risk measure optimisation model thatconsiders transactional or managerial costs in the objective function. This problem is essentially covered by Alexander et al. [6], where they formulate the CVaR optimal portfolio model. We shall use their work to form the optimal problem and solve it for our given portfolio(s).

(Obj. V). Chekhlov et al. [11] introduced a new one-parameter family of risk measures called Conditional Drawdown (CDD). We shall present the theorems and properties of this risk measure. We then formulate an optimal portfolio problem that shall be solved using the Conditional Drawdown (CDD) as either objective function or constraint. These problems generally lead to knapsack type optimisation problems and we shall use genetic algorithms to solve the problems.

The objectives highlighted above form the main research outcomes that we wish to achieve in this thesis. We shall present our investigations of each of the objectives above in a case study format where we present different scenarios which we aim to solve.

1.4. Remaining chapters

In Chapter 2, we give a general introduction to the different risk measures that have been developed and formed over the last 50 years. We also introduce some basic properties that make up the axiomatic framework for modern accepted risk measures and the fundamental theorems of the convexity which is needed in the optimisation framework for minimizing or maximizing risk measures. Some theoretical aspects of risk measure are presented for completeness as it is up to the reader to skip to the sub-section on summary of risk measures.

In Chapter 3, we focus on a select few risk measures such as distortion, spectral, coherent-distortion risk measures(CDRM) and drawdown risk measures. We explore the properties of each risk measure and how we can use these properties to form optimisation problems.

In Chapter 4, we introduce the fundamental concepts for general loss distributions and how these concepts are extended to form the optimisation problems for the CVaR risk measure. We formulate the CVaR optimisation problem and give methods for solving the CVaR optimisation problem efficiently.

In Chapter 5, we extend the theoretical work done in Section 3 and formulate the optimisation problems for the spectral, CDRM, and drawdown risk measures. We also introduce some computational methods for efficiently solving the alternative risk measure portfolio optimisation problems.

In Chapter 6, we take an interesting look at fat-tails and their effects on risk measures. We pay particular attention to how fat-tails affect CVaR as a risk measure.

In Chapter 7, we propose to solve various optimisation problems that either use risk measures as an objective function or a constraint. We present the optimisation problems as case studies. Further empirical analysis, results and discussions take place in this section.

Finally, in Chapter 8 we conclude the document by discussing major findings and a way to take the research further.

Figure 1-2 shows the section flow of how one may approach the reading of this document

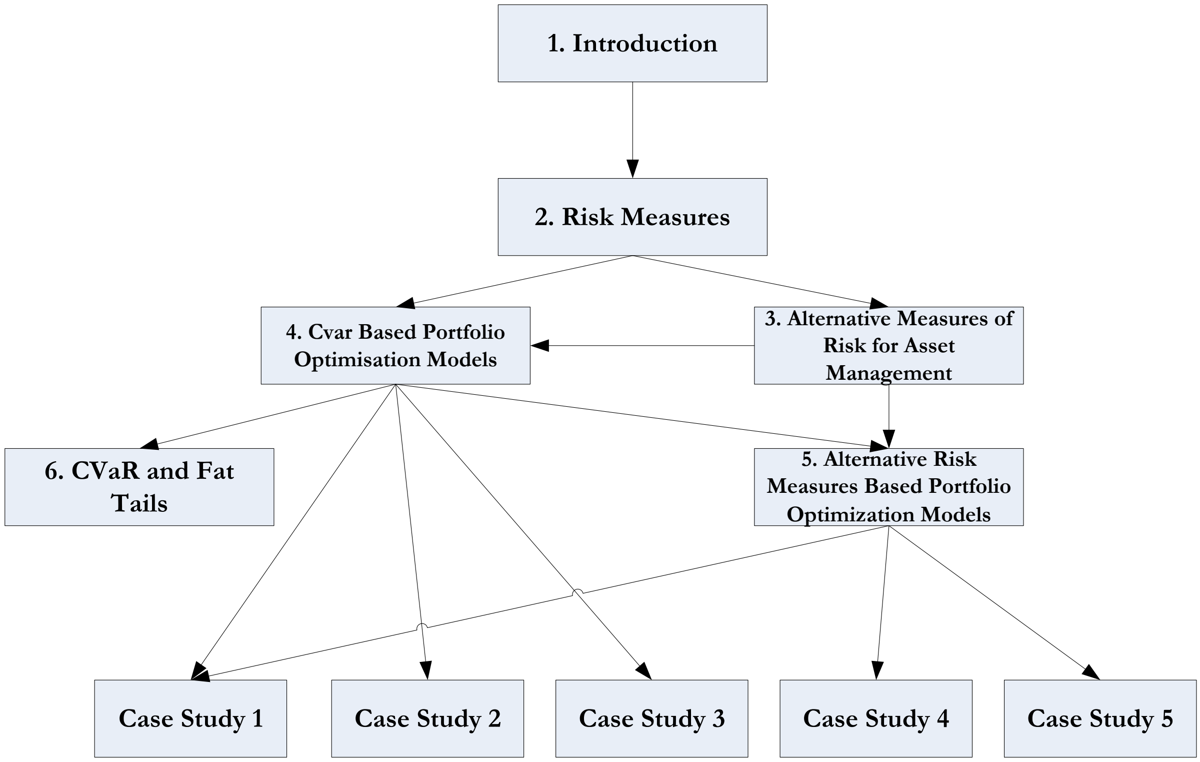


Figure 1-2: Section flow diagram

Based on the above flow diagram, the reader may move from Chapter 2 to either Chapter 3 or Chapter 4. If the reader so wishes to read CVaR based optimisation, then reading Chapter 2, 4 and 6 will be best suited. Case Studies that present work on CVaR optimisation is Case Study 1, 2 and 3.

If the reader wishes to read alternative risk measures, then Chapter 3 and 5 will be best suited. Case Studies 1, 4, and 5 cover alternative risk measure based optimisation.

Chapter 2. Risk measures

The purpose of this chapter is to present the various risk measures from a theoretical viewpoint. For applicability and practicality purposes portfolio returns over a given time horizon shall be considered. We shall present an axiomatic approach to risk management and develop an overview of risk measures. Much of the definitions and supporting theorems for risk measures shall be adopted from Artzner et al. [8].

2.1 Axiomatic approach to risk management

From a mathematical view, we shall consider a finite probability space (Ω, F, P) where Ω is the sample space of possible outcomes, F is the set of events and P assignment of probabilities to the events. Given a portfolio return X, we denote by F_X, the corresponding distribution function: $x \in \mathbb{R} \to F_X = P(X \leq x)$. A risk measure is a functional over the set of random portfolio returns mappings $X \to \rho(X) \in \mathbb{R}$. We can summarize the main axioms that may be fulfilled are, as applied in both theoretical and practical setting [8]:

(AX1.) Positive Homogeneity: for every random portfolio return X and real Value $\lambda > 0, \rho(\lambda X) = \lambda \rho(X)$.

(AX2.) Translation-invariance: for every random portfolio return X and real value $\alpha, \rho(X + \alpha) = \rho(X) - \alpha$.

(AX3.) Monotonicity: for every random portfolio return X and Y such that $X \geq Y$, $\rho(X) \leq \rho(Y)$.

(AX4.) Sub-additivity: for every random portfolio return X and Y, $\rho(X+Y) \leq \rho(X) + \rho(Y)$.

(AX5.) Law-invariance: for every random portfolio return X and Y with distribution functions F_X and F_Y, $F_X = F_Y \Rightarrow \rho(X) = \rho(Y)$.

(AX6.) Comonotonic: for every comonotonic random variable X and Y, $\rho(X+Y) = \rho(X) + \rho(Y)$.

Positive homogeneity (AX1.) also known as positive scalability signifies that a measure has the same dimension as a variable X. From a financial point of view positive homogeneity implies that a linear increase of the return leads to a linear increase of risk by the same positive factor.

Translational invariance (AX2.) implies that if the payoff increases by a known constant, the risk decreases with correspondence.

Monotonicity (AX3.) from a financial perspective implies that if one financial instrument payoff X is not less than another financial instrument payoff Y, then the risk of the X is not greater than the risk of the Y.

Sub-additivity (AX4.) forms the part of class of sums of risk and it states that the risk of a portfolio is not greater than the sum of the risks of the portfolio elements of components. In simpler terms, "a merger does not create extra risk" as quoted by Artzner et al. [8]. Law-invariance (AX5.) states that a risk measure does not depend on a risk itself but only on its underlying distribution.

2.2 Coherent measures of risk

In this subsection, the definition of risk measures shall be presented with some detail. This subsection and the next on convex risk measures are intended for completeness of the theoretical aspects and it is up to the reader to skip this subsection and proceed to the summary of risk measures subsection.

2.2.1 Definition of risk and coherence

Many initial papers defined and viewed risk in terms changes in values between two time instances. However, Artzner et al. [8] argues that risk is related to the variability of the future value of a position. Simply put, the future values of the position are more important and therefore should only be considered. Arztner et al. [8] put emphasis on understanding the nature of the future date and more importantly the measurement of risk of a position will be whether its future values belong to a subset of acceptable risks. Artzner et al. [8] considers three supervisors namely; a regulator, exchange clearing firm, and investment manager, each required weigh the trade-off between the severity of the risk measurement and the level of activities in the supervised domain.

Axioms on Acceptance Sets

The assumption is that the sets of all possible states of the world at the end of the period is known, but the probabilities of the different states may be unknown or not subjected to common agreement. In order to set the underlying axioms, the following notation shall be borrowed from Artzner et al. [8].

The notation that is adopted from Artzner et. al [8] and shall also be used in this section is:

i. Let Ω the set of states of nature, and assume it is finite. Considering Ω as the set of outcomes of an experiment, we compute the final net worth of a position for each element of Ω. It is a random variable denoted by X. Its negative part, $\max(-X,0)$ is denoted by X^- and the supremum of X^- is denoted by $\|X^-\|$. The random variable identically equal to 1 is denoted by 1. The indicator function of state ω is denoted by $1_{\{\omega\}}$

ii. Let G be the set of all risks, that is the set of all real valued functions on Ω. Since Ω is assumed to be finite, G can be identified with $\mathbb{R}^n$, where $n = card(\Omega)$. The cone of non-negative elements in G shall be denoted by L_+, its negative by L_-.

iii. Artzner et al. [8] call $A_{i,j}$, $j \in J_i$ a set of final net worths, expressed in currency i, which, in country i; are accepted by regulator/supervisor j:

iv. Artzner et al. [8] denote A_i the intersection $\bigcap_{j\in J_i} A_{i,j}$ and use the generic notation A in the listing of axioms below.

The following axioms shall now be presented from Artzner et al. [8].

Axiom 2.1

The acceptance set A contains L_+.

Axiom 2.2

The Acceptance set A does not intersect the set L_{--} where,

$$L_{--} = \{X \mid \text{for each } \omega \in \Omega, X(\omega) < 0\}. \qquad 2.1$$

Axiom 2.3

The acceptance set A satisfies $A \cap L_- = \{0\}$.

Axiom 2.4

The acceptance set is A convex.

Axiom 2.5

The acceptance set A is a positively homogeneous cone.

Artzner et al. [8] argue that all reasonable risk measures will have acceptance sets that satisfy Axioms 2.1 to 2.5.

Acceptance sets and measures of risk

Given some reference instrument, there is a natural way to define a measure of risk by defining how close or far from acceptance a position is.

Definition 2.1

A measure of risk is a mapping from G to $\mathbb{R}$.

Definition 2.2

Risk measure associated to an acceptance set. Given the total rate of return r on a reference instrument, the risk measure associated to the acceptance set A is the mapping from G to $\mathbb{R}$ denoted by $\rho_{A,r}$ and defined by

$$\rho_{A,r}(X) = \inf\left\{m \middle| m \cdot r + X \in A\right\}. \qquad 2.2$$

Definition 2.3

Acceptance set associated to a risk measure: the acceptance set associated to a risk measure ½ is the set denoted by A_ρ and defined by

$$A_\rho = \left\{X \in G \middle| \rho(X) \le 0\right\}. \qquad 2.3$$

Axiom 2.6

Translation invariance: for all $X \in G$ and all real numbers α, we have

$$\rho(X + \alpha \cdot r) = \rho(X) - \alpha\ . \qquad 2.4$$

Axiom 2.7

Subadditivity: for all X_1 and $X_2 \in G$ $\rho(X_1 + X_2) \le \rho(X_1) + \rho(X_2)$.

Axiom 2.7 reflects the idea that pooling risks helps to diversify a portfolio and decentralisation of a risk management. For example, if a risk manager has a total risk budget of B, he can devide B in to B_1 and B_2 where $B_1 + B_2 = B$. He can then allocate risk budget of B_1 and B_2 to different trading desks or operating units in the organization, safe in the knowledge that the firm-wide risk will not exceed B.

Axiom 2.8

Positive homogeneity: for all $\lambda \geq 0$ and all $X \in G$, $\rho(\lambda X) = \lambda \rho(X)$.

Axiom 2.8 reflects the fact that there are no diversification benefits when we hold multiples of the same portfolio.

Axiom 2.9

Monotonicity: for all X and $Y \in G$ with $X \leq Y$; we have $\rho(Y) \leq \rho(X)$.

Axiom 2.10

Relevance: for all $X \in G$ with $X \leq 0$ and $X \neq 0$; we have $\rho(X) > 0$.

Axiom 2.10 ensures that the risk measures identifies risky firm net values.

Definition 2.4

Coherence is a risk measure satisfying the four axioms of translation invariance (AX2.), subadditivity (AX4.), positive homogeneity (AX1.), and monotonicity (AX3.).

Axioms on acceptance sets and the axioms on measures of risks

It will be noticed that we claimed the acceptance set to be the fundamental object. We further discussed the axioms mostly in terms of the associated risk measure. The following propositions show that this was reasonable. For the proofs, please refer to Artzner et al. [8].

Proposition 2.1 [8]

If the set B satisfies Axioms 2.1, 2.2, 2.3 and 2.4, the risk measure $\rho_{B,r}$ is coherent. Moreover; $A_{\rho_{B,r}} = \overline{B}$ the closure of B.

Proposition 2.2 [8]

If a risk measure ρ is coherent, then the acceptance set A_ρ is closed and satisfies Axioms 2.1, 2.2, 2.3 and 2.4. Moreover $\rho = \rho_{A_p,r}$.

Proposition 2.3 [8]

If a set B satisfies Axioms 2.1, 2.2, 2.3 and 2.4, then the coherent risk measure $\rho_{B,r}$ satisfies the relevance axiom. If a coherent risk measure ρ satisfies the relevance axiom, then the acceptance set $A_{\rho_{B,r}}$ satisfies Axiom 2.2.

Artzner et al. [8] show in Proposition 2.2 that every coherent risk measure's acceptance set is closed and satisfies Axioms 2.1 to 2.5. Eq. (2.2) determines the minimal capital necessary to add to the firm to make the resulting firm's insolvency risk acceptable. Proposition 2.1 shows that if a acceptance set satisfies Axioms 2.1 to 2.5, thenthe risk meaure generated by the set is coherent and $A_{\rho_{B,r}} = \overline{B}$, the closure of B. Additionally, in Propostion 2.2 shows that the risk measure generated by a coherent measure's acceptance set is equal to the coherent measure, that is $\rho = \rho_{A_p,r}$.

2.3 Convex measures of risk

In this subsection, we continue to build on the theoretical foundations on the axioms presented in the preceding subsection and define and present the fundamental theorem and propositions for convex measures of risk as developed by Follmer and Schied [17, 18], and Follmer and Penner [16].

For Convexity, we relax the conditions of positive homogeneity and subadditivity we get the weaker property,

Convexity: $\rho\left(\lambda X+(1-\lambda)Y\right)\leq\lambda\rho(X)+(1-\lambda)\rho(Y)$ for any $\lambda\in[0,1]$. 2.5

Definition 2.5

A map $\rho: X\rightarrow\mathbb{R}$ will be called a convex measure of risk if it satisfies the conditions of convexity, monotonicity, and translation invariance.

Continuing on the definition of acceptance set associated to a risk measure we have the proposition below according to Foller and Schied [18].

Proposition 2.4

Suppose $\rho: X\rightarrow\mathbb{R}$ is a convex measure of risk with associated acceptance set A_ρ , Then $\rho_{A_\rho}=\rho$. Moreover, $A:=A_\rho$ enjoys the following properties A is convex and non-empty

v. if $X_1\in A$ and $Y\in X$ satisfies $Y\geq X_1$ then $Y\in A$,

v. if $X_1\in A$ and $Y\in X$, then $\left\{\lambda\in[0,1]\middle|\lambda X_1+(1-\lambda)Y\in A\right\}$ is closed in [0,1].

Proof Proposition 2.4 [18]

To show that $\rho_{A_\rho}=\rho$ for all X , note that the translation invariance of ρ implies that

$$\begin{aligned}\rho_{A,r}(X)&=\inf\left\{m\middle|m\cdot r+X\in A_\rho\right\}\\&=\inf\left\{m\middle|\rho(m+X)\leq 0\right\}\\&=\inf\left\{m\middle|\rho(X)\leq m\right\}\\&=\rho(X).\end{aligned} \qquad 2.6$$

The first two properties of $A := A_\rho$ are straightforward. The third one, we note that the function $\lambda \mapsto \rho(\lambda X + (1-\lambda)Y)$ is continuous as it is convex and takes only finite values. Hence, the set of $\lambda \in [0,1]$ such that $\rho(\lambda X + (1-\lambda)Y)$ is closed. □

Proposition 2.5 [18]

Assume that $A \neq 0$ is a convex subset of X which satisfies property 2 of Proposition 2.4, and denote by ρ_A the functional associated to A. If $\rho_A(0) > -\infty$, then

1. ρ_A is a convex measure of risk.
2. A is a subset of A_{ρ_A}. Moreover, if A satisfies property 3 of Proposition 2, then $A = A_{\rho_A}$.

For a detailed proof, see Follmer and Schied [17].

2.3.1 The representation theorem for convex risk measures

In this subsection, we present the proof of the structure theorem for convex measures of risk from the work of Follmer and Schied [17]. First, they considered the special case in which X is the space of the all real-valued functions on some finite set Ω. The representation theorem for convex risk measures forms the basis for the work that shall follow in the fourth coming sections. Hence the importance of the representation theorem and its proof is self-explaining. For further developments on this theorem and proof see Follmer and Schied [17] and the Appendix.

Theorem 2.1 [17]

Suppose X is the space of the all real-valued functions on a finite set Ω. Then $\rho : X \to \mathbb{R}$ is convex measure of risk if and only if there exists a "penalty function" $\alpha : \mathrm{P} \to (-\infty, \infty)$ such that

$$\rho(Z) = \sup_{Q \in P} \left(E_Q[-Z] - \alpha(Q) \right). \qquad 2.7$$

The function $\alpha(Q) \geq -\rho(0)$ satisfies for any $Q \in \mathrm{P}$, and it can be taken to be convex and lower semi-continuous on P.

Proof of Theorem 2.1 [17]

The "if" can be shown as follows: For each $Q \in \mathrm{P}$ the functional,

$$X \mapsto E_Q[-Z] - \alpha(Q) \quad 2.8$$

is convex, monotone, and translation invariant. These three properties are persevered under taking suprema.

Converse implication proof, we need the following auxiliary observation. For $Q \in \mathrm{P}$, define $\alpha(Q)$ by

$$\alpha(Q) = \sup_{X \in \mathrm{X}} \left(E_Q[-X] - \rho(X) \right). \quad 2.9$$

Then we claim that

$$\alpha(Q) = \sup_{X \in A_\rho} \left(E_Q[-X] \right) \quad 2.10$$

for the moment and denote the right hand side by $\widehat{\alpha}(Q)$. By definition of A_ρ we find $\alpha(Q) \geq \widehat{\alpha}(Q)$. To establish the converse inequality, take an arbitrary $X \in \mathrm{X}$ and recall that $X' := \rho(X) + X \in A_\rho$. Thus

$$\hat{\alpha}(Q) = \left(E_Q[-X] \right) = \left(E_Q[-X] - \rho(X) \right). \quad 2.11$$

This shows $\alpha(Q) = \widehat{\alpha}(Q)$. Note that we did not yet use the assumption that Ω is finite.

Now fix some $Y \in \mathrm{X}$ and take $\alpha(\cdot)$ as in Eq. (
2.11). Then we clearly have

$$\rho(Y) \geq \sup_{Q \in P} \left(E_Q[-Y] - \alpha(Q) \right). \quad 2.12$$

To establish the reverse inequality, take $m \in \mathbb{R}$ such that

$$m > \sup_{Q \in P} \left(E_Q[-Y] - \alpha(Q) \right). \quad 2.13$$

We must show that $m \geq \rho(Y)$ or, equivalently, $m + Y \in A_\rho$. Suppose on the contrary, $m + Y \notin A_\rho$. Since ρ is by definition a convex function on the Euclidean space $\mathbb{R}^\Omega$ taking only finite values, ρ is already continuous. Hence $A_\rho = \{\rho \leq 0\}$ is a closed convex set. Therefore, we can find linear functional ℓ on $\mathbb{R}^\Omega$ such that

$$\beta = \sup_{X \in A_\rho} \ell(X) < \ell(m + Y) = \gamma < \infty. \quad 2.14$$

It follows that ℓ is a negative linear functional. Indeed, note first that the axioms of normalization and monotonicity imply

$$\rho(X) \leq \rho(0) \text{ for } X \geq 0. \qquad 2.15$$

Thus, if $X \in \mathcal{X}$ satisfies $X \geq 0$, then $\lambda X + \rho(0) \in A_\rho$ for all $\lambda \geq 1$, and hence

$$\gamma > \ell(\lambda X + \rho(0)) = \lambda \ell(X) + \rho(0). \qquad 2.16$$

Taking $\lambda \uparrow \infty$ yields that $\ell(X) \leq 0$. If we assume that ℓ applied to the constant function 1 gives -1, what we can do without loss of generality, then

$$Q[A] := \ell(-I_A) \qquad 2.17$$

defines a probability measure $Q \in \mathcal{P}$. By Eq. (2.12) and Eq. (2.14) we find

$$\alpha(Q) = \sup_{X \in A_\rho} \left(E_Q[-X]\right) = \beta. \qquad 2.18$$

But

$$E_Q[-Y] - m = \ell(m+Y) = \gamma > \beta = \alpha(Q) \qquad 2.19$$

which is a contradiction to our choice of m. Therefore, we must have $m + Y \in A_\rho$ and, thus $m \in \rho(Y)$. □

2.4 Summary of risk measures

As shown above, a risk measure, $X \to \rho(X) \in \mathbb{R}$ is functional that assigns a numeric value to a random variable representing a random return or payoff. Not every functional corresponds to the intuitive notion of risk. One of the main characteristics of such a function is that a higher uncertain return should conform to a higher functional value. We shall now present some of the various risk measures that were theoretically developed over time.

2.4.1 Pederson and Satchell's

Pederson and Satchell (see Rachev et al. [32] for details) defined their own class of risks that is seen as a deviation from a location measure. Nonnegativity, positive homogeneity, sub-additivity, and translation invariance are considered desirable properties of a "good financial risk measure". Pedersen and Satchell presented in their work the full characterization of the appropriate risk measures that was based on their system of axioms.

2.4.2 Coherent risk measures

Coherent risk measures were introduced by Artzner et al. [8]. Coherent risk measures are those measures which are translation invariant, monotonous, sub-additive, and positively homogeneous as defined above. Coherent measures have the following general form:

$$\rho(X) = \sup_{Q \in \wp} E_Q[-X] \qquad 2.20$$

where $\wp$ is some class of probability measures on Ω . Four criteria proposed by Artzner et al. [8] provide rules for selecting and evaluating risk measures. These axioms have been detailed earlier in this chapter. Note that one should be aware that not all risk measures satisfying the four proposed axioms are reasonable to use under certain practical situations. Wang et al. [40] suggested that "a risk measure should go beyond coherence" in order to utilize useful information in a large part of a loss distribution. Dhaene et al. (see Rachev et al. [32] for further details), observing "best practice" rules in insurance, concluded that coherent risk measures "lead to problems".

2.4.3 Convex risk measures

Convex risk measures were studied by Follmer and Schied [16, 17, 18] and Frittelli and Rosazza Gianin (see Rachev at al. [32] for further details). The generalization of coherent risk measures derived by relaxation of the positive homogeneity assumption, together with the sub-additivity condition leads to the basis for convexity. Any convex risk measure takes into account a nonlinear increase of the risk with the size of the position and has the following structure [32]:

$$\rho(X) = \sup_{Q \in \wp}\left(E_Q[-X] - \alpha(Q)\right) \qquad 2.21$$

where α is a penalty function defined on probability measures on Ω. The representation theorem and other important features pertaining to convex risk measure have been presented earlier in this chapter. Further, the requirements for convex risk measures in the CVaR optimisation based model will serve great advantages.

2.4.4 Law invariant coherent risk measures

From the notation of Kusuoka (see Rachev et al. [32] for further details), law invariant coherent risk measures have the form:

$$\rho_\alpha(X) \overset{\Delta}{=} \frac{1}{\alpha}\int_{1-\alpha}^{1} Z_{-X}(x)\,dx\,, \qquad 2.22$$

where $Z:[0,1) \to \mathbb{R}$ is non-decreasing and right continuous. This class of risk measures satisfies the lower semi-continuity property for all $X \in L^\infty, 0 \le \alpha \le 1$. The class of insurance prices characterized by Wang et al. [40] is an example of law invariant coherent risk measures.

2.4.5 Spectral risk measures

Spectral measures of risk add two axioms to the set of coherency axioms: law invariance (AX5.) and comonotonic additivity (AX6.). Spectral risk measures consist of a weighted average of the quantiles of the returns distributions. The spectrum denoted by ϕ is a non-increasing weight function of these weighted quantiles of a return distribution. Spectral risk measure is defined as follows:

$$M_\phi(X) = -\int_0^1 \phi(x) F_X(x)\,dx\,. \qquad 2.23$$

Where ϕ is a non-negative, non-increasing, right-continuous integrable function defined on

[0, 1] and such that $\int_0^1 \phi(x)\,dx = 1$.

The coherency of the spectral risk measure is dependent on the assumptions made on ϕ.

If any of these assumptions are relaxed, the measure is no longer coherent. Additionally, spectral risk measures possess consistency with second order stochastic dominance (SSD), and expected utility theory. More of the details around spectral risk measure shall be presented in the next chapter.

2.4.6 Deviation measures and expectation-bounded risk measures

Rockafeller et al. [35, 36] defined deviation measures as:

i. positive,
ii. sub-additive (AX4.),
iii. positively homogeneous (AX1.),
iv. Gaivoronsky-Pflug (G-P) translation- To get a detail account on this property refer to the appendix in Rockafeller et al. [35].

Deviation measures are typically useful to the totally risk-averse investors. Rockafeller et al. [35, 36] proposed expectation-bounded risk measures, imposing the conditions of:

i. sub-additivity (AX4.),
ii. positive homogeneity (AX1.),
iii. translation invariance (AX2.) and
iv. additional property of expectation-boundedness.

There is a correspondence in the one-to-one relationship between deviation measures and expectation-bounded risk measures. One can derive expectation-bounded coherent risk measures if additionally, monotonicity is satisfied [32].

2.4.7 Parametric classes of risk measures

Stone (see Rachev et al. [32] for further details) defined a general three-parameter class of risk measures, which has the form,

$$R\left[c,k,A\right]=\left(\int_{A}^{-\infty}\left|y-c\right|^{k}f(y)dy\right)^{1/k} \qquad 2.24$$

where A, $c\in\mathbb{R}$ and $k>0$. Stone's class of risk measures includes several commonly used measures of risk and dispersion. These measures of risk and dispersion that is used are the standard deviation, the semi-standard deviation, and the mean absolute deviation.

Pedersen and Satchell (see Rachev et al. [32] for further details) generalized Stone's class of risk measures. They introduced the five-parameter class of risk measures, defined as,

$$R\left[A,c,\alpha,\theta,\omega\right]=\left(\int_{A}^{-\infty}\left|y-c\right|^{\alpha}\omega\left[F\left(y\right)\right]f(y)dy\right)^{\theta}, \qquad 2.25$$

for a considered bounded function $\omega(\cdot)$, A, $c\in\mathbb{R}$, $\alpha>0$, $\theta>0$.

2.4.8 Quantile-based risk measures

Quantile-based risk measures include Value-at-Risk (VaR), expected shortfall (ES), tail conditional expectation (TCE), and worst conditional expectation (WCE). We shall introduce each measure below briefly; with further details the reader is referred to Rachev et al. [32], Einmal et al. [13], Pflug [29], Mausser and Rosen [26], Mckay and Keefer [27]. Value-at-Risk (VaR) specifies how much one can lose with a given probability. It is defined as

$$VaR_\alpha(X) = -x^{(\alpha)} = q_{1-\alpha}(-X). \quad 2.26$$

VaR has the following properties:

i. Monotonicity (AX3.),
ii. positive homogeneity (AX1.),
iii. translation invariance (AX2.),
iv. law invariance (AX5.),
v. comonotonic additivity (AX6.).

Expected shortfall (ES), also known as tail (or conditional) VaR, corresponds to the average of all VaR_α's above the threshold α,

$$ES_\alpha(X) = \frac{1}{1-\alpha}\int_\alpha^1 VaR_\upsilon(X) d\upsilon. \quad 2.27$$

ES was proposed in order to overcome some of the theoretical weaknesses of VaR. ES has the following properties:

i. law invariance (AX5.),
ii. translation invariance (AX2.),
iii. comonotonic additive (AX6.),
iv. continuity,
v. monotonicity (AX3.),
vi. sub-additivy (AX4.).

Tail conditional expectation (TCE) was proposed by Artzner et al. [8] in the following form:

$$TCE_\alpha(X) = -E\left\{X \mid X \le x^{(\alpha)}\right\}. \quad 2.28$$

TCE_α does not possess the sub-additivity property for general distributions. It further has coherency for continuous distributions only. Worst conditional expectation (WCE) is defined as

$$WCE_\alpha(X) = -\inf\left\{E[X \mid B] : B \in A, P(B) > \alpha\right\}. \quad 2.29$$

$WCE_\alpha(X)$ is not law-invariant, so it cannot be estimated solely from data. Using such measures can lead to different risk values for two portfolios with identical loss distributions. This forms a practical weakness and as such is rarely used.

A comparison of ES, TCE, and WCE, yields the following relationship according to Rachev et al. [32]:

$$TCE_\alpha(X) \le WCE_\alpha(X) \le ES_\alpha(X). \quad 2.30$$

When the underlying probability law varies, ES has the maximum value among TCE and WCE. If the distribution of X is continuous, then the following can be realised,

$$TCE_\alpha(X) = WCE_\alpha(X) = ES_\alpha(X). \quad 2.31$$

2.4.9 Drawdown measures

Drawdown measures are intuitive risk measures. Drawdowns measure the difference between two observable quantities. These observable quantities can be local maximum and local minimum of the portfolio wealth. Chekhlov et al. [11] defined the drawdown function as the difference between the maximum of the total portfolio return up to a specific time t and the portfolio value at t. Drawdown measures can be compared to the notion of deviation measures. Examples of drawdown measures constitute, see later Chapter 3 that give detail definitions for each:

i. Absolute drawdown (AD),
ii. maximum drawdown (MDD),
iii. Average drawdown (AvDD),
iv. Drawdown at risk (DaR), and
v. Conditional drawdown at risk (CDaR).

Drawdown measures have computational simplicity, and therefore drawdown measures cannot describe the real situation on the market, and therefore, should be used in combination with other measures.

2.4.10 Distortion risk measures

A distortion risk measure can be defined as the: "distorted expectation of any non-negative loss random variable X" [32]. It is accomplished by using the distortion function g as follows,

$$\rho_g(X) = \int_0^\infty g\left(1 - F_X(x)\right)dx = \int_0^1 F_X^{-1}(x)(1-q)\,dg(q). \quad 2.32$$

Where $g:[0,1] \to [0,1]$ is a continuous increasing function with $g(0)=0$ and $g(1)=1$;

$F_X(x)$ denotes the cumulative distribution function of X, while $g\left(F_X(x)\right)$ is referred to as a distorted distribution function.

For the profit and loss distributions, when the loss random variable can take any real number, the distortion risk measure is:

$$\rho_g(X)=\int_0^1 F_X^{-1}(x)\,dH(x)=-\int_{-\infty}^0 H(F_X(x))\,dx+\int_0^\infty\left[1-H(F_X(x))\right]dx \qquad 2.33$$

where $H(u)=1-g(1-u)$. A similar expression holds if we use the survival function $S_X(x)=1-F_X(x)=P(X>x)$ instead of the distribution function,

$$\rho_g(X)=-\int_{-\infty}^0\left[1-g(S_X(x))\right]dx+\int_0^\infty g(S_X(x))\,dx\,. \qquad 2.34$$

Further details of the distortion risk measure shall be presented in the next chapter.

2.5 Remarks

In Chapter 2, we presented the axiomatic framework that defines the various risk measures that have been devloped over the past decades. We have given a detail account of the relevant axioms for coherency and have shown the representation theorems for convex risk measures. We have also given broad summary of the various risk measures that have been developed. This chapter's aim was to achieve partial objectives of Obj. (I) and Obj. (V).

Chapter 3. Alternative measures of risk for asset management

In this chapter, we explore into the analysis of the different risk measures that were introduced in the preceding chapter. We shall focus on the alternative risk measures to the VaR and CVaR that have been researched in the recent years. The main question for asset manager is the choice of an adequate risk measure. The answer is not obvious as it is generally not easy to identity which particular risk measure might be best suited. Further difficultly arises as there is no clear way of comparing risk measures and there is no guarantee that the risk measure will be "good" under all circumstances.

We shall introduce details pertaining to the spectral, coherent-distortion and distortion risk measures. Some key theorems shall be presented that serve as building blocks that allow one to use these alternative risk measures for portfolio optimisation.

3.1 Distortion risk measures

A distortion risk measure can be defined as the: "distorted expectation of any non-negative loss random variable X" [32]. Rachev et al. [32] show based on using a "dual utility" or the distortion function g that

$$\rho_g(X) = \int_0^\infty g\left(1 - F_X(x)\right)dx = \int_0^1 F_X^{-1}(x)(1-q)\,dg(q)\,. \qquad 3.1$$

Where $g:[0,1]\to[0,1]$ is a continuous increasing function with $g(0)=0$ and $g(1)=1$; $F_X(x)$ denotes the cumulative distribution function of X, while $g\left(F_X(x)\right)$ is referred to as a distorted distribution function.

For the Profit/Loss-distributions, when the loss random variable can take any real number, the distortion risk measure is obtained as follows:

$$\rho_g(X) = \int_0^1 F_X^{-1}(x)\,dH(x) = -\int_{-\infty}^0 H\left(F_X(x)\right)dx + \int_0^\infty \left[1 - H\left(F_X(x)\right)\right]dx \qquad 3.2$$

where $H(u) = 1 - g(1-u)$. We achieve a similar expression if we use the survival function $S_X(x) = 1 - F_X(x) = P(X > x)$ instead of the distribution function,

$$\rho_g(X) = -\int_{-\infty}^0 \left[1 - g\left(S_X(x)\right)\right]dx + \int_0^\infty g\left(S_X(x)\right)dx\,. \qquad 3.3$$

A more general class of distortion risk measures was developed, depending on the choice of parameters α , g, and h. It has the following form,

$$H_{\alpha,g,h}(X)=\alpha+H_h\left((X-\alpha)^+\right)-H_g\left((\alpha-X)^+\right) \quad 3.4$$

where $\alpha^+=\max[0,\alpha]$. When $\alpha=0$ and $H(x)=1-g(1-x)$, then we again obtain the Choquet integral representation.

3.1.1 Distortion risk measure properties

We shall introduce some properties of distortion risk measures that are derived and correspond to the Choquet Integral. The reader is advised to refer to Rachev et al. [32] for details regarding these properties.

i. If $X\geq 0$, then $\rho_g(X)\geq 0$ monotonicity.

ii. $\rho_g(\lambda X)=\lambda\rho_g(X)$, for all $\lambda\geq 0$, positive homogeneity.

iii. $\rho_g(X+c)=\rho_g(X)+c$, for all %, translation invariance.

iv. $\rho_g(-X)=-\rho_{\tilde{g}}(X)$, where $\tilde{g}(x)=1-g(1-x)$.

v. If a random variable X_n has finite number of values and $\rho_g(X)$ exists, then $\rho_g(X_n)\rightarrow\rho_g(X)$.

vi. If X and Y are comonotonic risks, taking positive and negative values, then $\rho_g(X+Y)=\rho_g(X)+\rho_g(Y)$, this is called comonotonic addivity.

vii. In the generalized case, distortion risk measures are not additive $\rho_g(X+Y)\neq\rho_g(X)+\rho_g(Y)$.

viii. Distortion risk measures are sub-additive if and only if the distortion function g is concave. This can be represented as follows $\rho_g(X+Y)\leq\rho_g(X)+\rho_g(Y)$.

ix. For a non-decreasing distortion function g , the associated risk measure ρ_g is consistent with the stochastic dominance of order 1, $X\leq_1 Y\Rightarrow\rho_g(X)\leq\rho_g(Y)$.

x. For a non-decreasing concave distortion function g , the associated risk measure g is consistent with the stochastic dominance of order 2, $X\leq_2 Y\Rightarrow\rho_g(X)\leq\rho_g(Y)$.

xi. For a strict concave distortion function g , the associated risk measure g is strictly consistent with stochastic dominance of order 2, $X<_2 Y\Rightarrow\rho_g(X)<\rho_g(Y)$.

3.1.2 Examples of distortion risk measures

As shown above, the distortion risk measure depends on the distortion risk function. Challenges come when the "best" distorted risk measure is needed to find the "best" suited distorted risk function. From Rachev et al. [32], we present some well-known distortion risk functions below.

i. With $g(x)=x$, we have $\rho_g(X)=E\left[X\right]$, assuming the mathematical expectation exists.

ii. VaR corresponds to the distortion function:

$$g(x)=\begin{cases} 0, & \text{if } x<1-p, \\ 1, & \text{if } x\geq 1-p. \end{cases} \tag{3.5}$$

iii. CVaR can be defined as distortion risk measure based on the distortion function: $g(x)=\min\left(\frac{x}{1-p},1\right), \quad x\in\left[0,1\right]$. It is interesting to note the CVaR s not differentiable at $x=1-p$. Due to this, it discards potentially valuable information since all mapping of percentiles below $x=1-p$ to a single point "0" is done. This implies it does not take into account the severity of extreme values or events.

iv. In order to overcome this the following distortion functions shall be considered: $g(x)=\Phi\left(\Phi^{-1}(x)-\Phi^{-1}(q)\right)$ for $p\in\left[0,1\right]$, where $0<q\leq 0.5$ is some parameter. The distortion function g is non-decreasing, concave, and such that $g(0)=0$ and $g(1)=1$. This risk measure corresponds to the well-known Wang Transform.

v. The beta family of distortion risk measures is given as, with incomplete beta function:

$$g\left(F_X(X)\right)=\beta\left(a,b;F_X(X)\right)=\int_0^{F_X(x)}\frac{1}{\beta(a,b)}t^{b-1}\,dt=S_\beta\left(F_X(x)\right). \tag{3.6}$$

Where $S_\beta(x)$ is the distribution function of the beta distribution, and $\beta(a,b)$ is the beta function with parameters $a>0$ and $b>0$, that is

$$\beta(a,b)=\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}=\int_0^1 t^{a-1}(1-t)^{b-1}\,dt\,. \tag{3.7}$$

Beta functions are concave if and only if when $a\leq 1$ and $b\geq 1$; and strictly concave if a and b are both not equal to 1.

vi. The proportional Hazard (PH) transforms is a special case of the beta-distortion risk measure with $a=0.1, b=1$. The PH-transform risk measure is defined as,

$$\rho_{PH}(X)=\int_0^\infty S_X(x)^{\frac{1}{\lambda}}\,dx, \quad \lambda>1 \tag{3.8}$$

where $S_X(x) = 1 - F_X(x)$.

3.2 Spectral risk measures

Spectral measures of risk can be defined by adding law invariance and comonotonic additivity axioms to the set of coherency axioms. Spectral risk measures consist of a weighted average of the quantiles of the returns distribution using a non-increasing weight function. This function can be referred to as a spectrum and is denoted by ϕ. It is defined as follows:

$$M_\phi(X) = -\int_0^1 \phi(x) F_X(x)\, dx . \quad 3.9$$

Where ϕ is a non-negative, non-increasing, right-continuous integrable function defined on [0, 1] and such that $\int_0^1 \phi(x)\, dx = 1$. Coherency of spectral risk measures depends on assumptions made on ϕ. If any of these assumptions are relaxed, the measure is no longer coherent. Spectral risk measures possess:

i. positive homogeneity,
ii. translation invariance,
iii. monotonicity,
iv. sub-additivity,
v. law invariance,
vi. comonotonic additivity,
vii. consistency with second order stochastic dominance (SSD) and expected utility theory.

The coherence of spectral risk measures comes from the assumption made on the spectrum. Thus if the spectrum assumptions change then the measure is no longer the same. From Adam et al. [3] we see that spectral risk measure can be expressed as an empirical distribution of portfolio returns as:

$$M_\phi(X) = -\sum_{i=1}^{n} \lambda_i x_i , \quad 3.10$$

where $\lambda_i = \int_{(i-1)/n}^{i/n} \phi(p)dp \geq 0, \sum_{i=1}^{n} \lambda_i = 1$. The computation of any spectral risk measure involves an average value of ranked or sorted portfolio returns. Due to this property, expected shortfall is a special spectral risk measure with $\phi(p) = \frac{1}{\alpha} \times 1_{[0,\alpha]}(p)$and it can be shown that any spectral risk measure can be expressed as a weighted average of expected shortfalls.

3.2.1 Generation of spectral risk measures from expected shortfall [1]

Acerbi [1] view the expected shortfall as, let $F_X(x) = P[X \le x]$ be the distribution function of the profit–loss X of a given portfolio Π and define the usual generalized inverse of $F_X(x)$ as

$$F^{\leftarrow}{}_X(p) = \inf\{x \mid F_X(x) \ge p\}. \qquad 3.11$$

The α -expected shortfall defined for $\alpha \in (0,1]$ as

$$ES_{(\alpha)}(X) = -\frac{1}{\alpha}\int_0^{\alpha} F^{\leftarrow}{}_X(p)dp\ , \qquad 3.12$$

can be shown by Acerbi et al. [21], to be a risk measure satisfying the axioms of Definition 2.1. For $\alpha = 0$ it is natural to extend the definition of $ES_{(0)}(X)$ as the very worst case scenario,

$$ES_{(0)}(X) = -ess.\inf\{X\}. \qquad 3.13$$

Recall that the expected shortfall is closely related but not coincident to the notion of conditional value at risk $CVaR_\alpha$ or tail conditional expectation TCE_α defined as [1]

$$CVaR_\alpha(X) = TCE_\alpha(X) = -E\left[X \mid X \le F_X^{\leftarrow}(\alpha)\right]. \qquad 3.14$$

In fact, conditional Value-at-Risk is not a coherent measure in general. It coincides with ES_α (and it is therefore coherent) only under suitable conditions such as the continuity of the probability distribution function $F_X(x)$ [12].

Introducing a measure $d\mu(\alpha)$ on $\alpha \in [0,1]$, and under suitable integrability conditions, we get:

$$M_\mu(X) = \int_0^1 d\mu(\alpha)\alpha ES_\alpha(X) = -\int_0^1 d\mu(\alpha)\int_0^\alpha dpF_X^{\leftarrow}(p) \qquad 3.15$$

is a risk measure as long as the normalization condition,

$$\int_0^1 \alpha\, d\mu(\alpha) = 1 \qquad 3.16$$

is satisfied. Based on the Fubini–Tonelli theorem (see Acerbi [1] for details) we can interchange the integrals,

$$M_\mu(X) = -\int_0^1 dpF_x^{\leftarrow}(p)\int_p^1 d\mu(\alpha) \equiv -\int_0^1 dpF_x^{\leftarrow}(p)\phi(p) \equiv M_\phi(X). \qquad 3.17$$

It is easy to see that the parametrization in terms of any measure $d\mu(\alpha)$ can be traded with a parametrization in terms of a decreasing positive risk spectrum as $\phi(p)=\int_1^p d\mu(\alpha)$.The normalization condition Eq. (3.17) translates into the following normalization condition for ϕ,

$$\int_0^1 \phi(p)dp = \int_0^1 dp \int_p^1 d\mu(\alpha) = \int_0^1 d\mu(\alpha)\int_0^\alpha dp = \int_0^1 d\mu(\alpha)\alpha = 1 . \qquad 3.18$$

3.2.2 Estimation of spectral risk measures

The risk measure M_φ is a very simple object to be used in practice. The integral of Eq. (3.9) is computable only when an explicit analytical expression for the inverse cumulative distribution function $F_x^{\leftarrow}(p)$ is available. In a practical risk management system this is typically the case only if the approach chosen for the probability distributions is parametric.

Acerbi [1] suggest the best method for evaluating M_ϕ is not by its integral definition, but rather by the estimator $M_\phi^{(N)}$ on a sample of N i.i.d. realizations $X_1,\ldots,X_N$ of the portfolio profit–loss X. They define it by introducing the ordered statistics $X_{i:N}$ given by the ordered values of the N-tuple $X_1,\ldots,X_N$. In other words: $\{X_{1:N},\ldots,X_{N:N}\}=\{X_1,\ldots,X_N\}$ and $X_{1:N} \le X_{2:N} \ldots \le X_{N:N}$. Following the work of Acerbi [1], definitions and theorems will be kept in accordance with their work.

Definition 3.1 [1]

Let $X_1,\ldots,X_N$ be N realizations of an r.v. X. For any given N-tuple of weights $\phi_{i=1,\ldots,N} \in \mathbb{R}$ we define the statistics

$$M_\phi^{(N)}(X) = -\sum_{i=1}^{N} X_{i:N}\phi_i . \qquad 3.19$$

Note that $M_\phi^{(N)}$ the spectral risk measure generated by ϕ_i.

Definition 3.2 [1]

An N-tuple $\phi_{i=1,\ldots,N} \in \mathbb{R}$ is said to be an “admissible” risk spectrum if [1]

i. ϕ_i is positive, $\phi_i \ge 0$,

ii. ϕ_i is decreasing, $\phi_i \ge \phi_j$ if $i < j$,

iii. $\sum_i \phi_i = 1$.

Theorem 3.1 [1]

The spectral risk measure $M_{\phi}^{(N)}$ of Definition 3.1 is a risk measure for any fixed $N \in \mathbb{N}$ if and only if ϕ_i is an admissible risk spectrum.

Theorem 3.1 has a wide range of applications. It provides a risk measure for a sample of N realizations of a random variable X. Since the theorem holds for any finite $N \in \mathbb{N}$, the coherency of the measure is not related to some law of large numbers. This result is immediately applicable in any scenario-based risk management system.

In a practical setup, an investor should choose his/her own risk averse function ϕ_i to assess her risks independently of the number of scenarios available for the estimation of M_{ϕ}. We can consider $\phi(p)$ as a positive decreasing normalized function rather than an abstract element of $L^1([0,1])$. Given $\phi(p)$ and fixed a number N of scenarios, the most natural choice for an admissible sequence i is given by

$$\phi_i = \frac{\phi\left(i/N\right)}{\sum_{k=1}^{N}\phi\left(k/N\right)}, \qquad i = 1,\ldots,N \quad . \qquad 3.20$$

The above expression satisfies $\sum_i \phi_i = 1$ for any finite N. The investor can then use the spectral risk measure $M_{\phi}^{(N)}$ generated by this sequence as a risk measure based on Theorem 3.1, it ensures its coherence for any finite N.

We have then shown that $M_{\phi}^{(N)}$ provides not only a coherent measure for any fixed N, but also a consistent way for estimating, for large number of scenarios the risk measure M_{ϕ}.

3.3 Coherent- Distortion risk measures

There are two ways to derive and define coherent distortion risk measure (CDRM) according to Feng and Tan [15]. Feng and Tan [15] define the CDRM as a subclass of distortion risk measure (DRM), namely DRM with concave distortion function g; they also defined CDRM as a subclass of CRM, namely CRM that is also comonotonic and law invariant. These two definitions are indeed equivalence since it is shown in Feng and Tan [15] that the class of coherent distortion risk measures coincides with the class of comonotonic law invariant coherent risk measures. Following the work of Feng and Tan [15], definitions and theorems will be kept in accordance with their work.

Definition 3.3 [15]

We say ρ is a coherent distortion risk measure (CDRM) if,

i. ρ_g is a distortion risk measure (DRM) with a concave distortion function g, or equivalently,
ii. ρ is a coherent risk measure (CRM) that is also comonotonic and law-invariant.

The following representation theorem for CDRM is the key result that enables us to develop a convex optimisation framework for any CDRM portfolio selection problem.

Theorem 3.2 [15]

For any random variable X and a given concave distortion function g, risk measure ρ_g is a CDRM if and only if there exists a function (coherent-distortion function) $w:[0,1]\mapsto[0,1]$, satisfying $\int_{\alpha=0}^{1} w(\alpha)d\alpha = 1$, such that,

$$\rho_g(X)=\int_{\alpha=0}^{1} w(\alpha)\phi_\alpha d\alpha \qquad 3.21$$

where $\phi_\alpha(X)$ is the α-CVaR of X. This representation theorem says that any CDRM can be represented as a convex combination of $CVaR_\alpha(X)$, $\alpha\in[0,1]$ and based in this one can construct any CDRM based on a convex combination of $CVaR_\alpha(X)$. This result was proved by Feng and Tan [15] for continuous portfolio loss distributions. Feng and Tan [15] proved and strengthened the representation theorem that any CDRM can be represented as a convex combination of finite number of $CVaR_\alpha(X)$ based on the assumption that the portfolio loss has discrete uniform distribution.

Definition 3.4 [15]

For a given loss observation $L=(l_1,...,l_m)$ and the corresponding ordered losses $l_1<l_2<...<l_m$. Let $p_{(i)}$ be the probability of realizing l_i, $i=1,...,m$ and let $S_i(l_i)=1-\sum_{j=1}^{i} p_{(i)}$. Define a

CVaR-matrix $Q \in \mathbb{R}^m \times \mathbb{R}^m$ with columns $Q_i \in \mathbb{R}^m$, i=1,...,m as

$$Q = [Q_1, Q_2, ..., Q_m]$$
$$= \begin{bmatrix} p_1 & 0 & 0 & \cdots & 0 \\ p_2 & \frac{p_2}{1-S_l(l_1)} & 0 & \cdots & 0 \\ p_3 & \frac{p_3}{1-S_l(l_1)} & \frac{p_3}{1-S_l(l_2)} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_m & \frac{p_m}{1-S_l(l_1)} & \frac{p_m}{1-S_l(l_2)} & \cdots & \frac{p_m}{1-S_l(l_{m-1})} = 1 \end{bmatrix}. \quad 3.22$$

Since portfolio losses are discretely distributed at m points, there are m jumps in the cumulative function of l_i, $i = 1,...,m$. By defining

$$\alpha_i = \begin{cases} 0 & for\ i = 1 \\ \sum_{j=1}^{i-1} p(j) & for\ i = 2,...,m \end{cases} \quad 3.23$$

at these m jumps, then the m CVaR's at these probability levels are then given by

$$\varphi_{\alpha_i}(l) = \frac{1}{1-\alpha_i}\sum_{j=1}^{m} p(j) l_j = \sum_{j=1}^{m} \frac{p(j)}{1-S_l(l_{m-1})} l_j = \sum_{j=1}^{m} Q_{ij} l_j, \quad 3.24$$

for $i = 1,..,m$ and Q_{ij} is the (i; j)-th entry of Q. Note that column Q_i is essential to the calculation of $CVaR_{(i-1)/m}(l)$.

Feng and Tan [15] give a finite generation result for the CDRM which shall be presented in Theorem 3.3 below. Theorem 3.3 is generalized form of Theorem 3.2 applied to general discrete loss distributions.

Theorem 3.3 [15]

For a give portfolio loss sample $L = (l_1,...,l_m)$, the corresponding ordered losses $(l_1,...,l_m)$ and a given concave distortion function g, the resulting CDRM ρ_g is given by,

$$\rho_g(l) = \sum_{i=1}^{m} q_i l_i \quad 3.25$$

where q_i ,$i = 1,...,m$ are defined in Eq. (3.23). Moreover, every such **q** can be written in the form

$$q = Qw, \quad 3.26$$

where $w^T = (w_{1,},...,w_m)$ denotes the convex weights satisfying $w_i \geq 0,\ i = 1,...,m$ and $\sum_{i=1}^{m} w_i = 1$, and **Q** is to the CVaR-matrix. The convex weights **w** are given by

$$w_i = \begin{cases} \dfrac{q_1}{p_1}, & \text{if } i = 1, \\ \left(q_i - \dfrac{p_i}{p_{i-1}} q_{i-1}\right) \dfrac{S_l(l_{i-1})}{p_i}, & \text{if } i = 2,..,m. \end{cases} \qquad 3.27$$

The following observations can be made. First, it is easy to verify that the convex weights defined in Eq. (3.27) satisfy $w_i \geq 0,\ i = 1,...,m$ and $\sum_{i=1}^{m} w_i = 1$. Theorem 3.3 implies that every CDRM can be defined as a convex combination of the ordered losses $1(l_1,...,l_m)$ via Eq. (3.25) or equivalently as a convex combination of CVaR's via Eq. (3.26). The latter formulation is what Feng and Tan [15] adopt in their CDRM portfolio optimisation model.

3.4 Drawdown risk measures

In this subsection, we present some general properties and theorems regarding the absolute drawdown for a single sample path, maximum, average and conditional drawdowns, and multi-scenario drawdown measure. Following the work of Chekhlov et al. [11], definitions and theorems will be kept in accordance with their work.

3.4.1 Absolute drawdown for a single sample path

We present the notion of the Absolute Drawdown (AD). The AD is applied to a sample path of the uncompounded cumulative portfolio rate of return. Note that the AD is applied not to the compounded cumulative portfolio rate of return $W_k(x(t_k))$. If the values of $r_k^p(x(t_k))$ for $k = 1,..., N$ determine a sample path of the portfolio's rate of return, then, the uncompounded cumulative portfolio rate of return at time moment t_k is given by:

$$w_k(x(t_k)) = \begin{cases} 0, & \text{if } k = 0, \\ \sum_{l=1}^{k} r_l^p(x(t_l)), & \text{if } k = 1,...,N. \end{cases} \qquad 3.28$$

We use w_k instead of $w_k(x(t_k))$, as w_k is always a function of vector $x(t_k)$. In this subsection, we shall consider only a single sample path of w_k , $k = 1,..., N$, which will be denoted by vector w.

Definition 3.5 [11]

The AD is a vector-variable functional depending on the sample path w as is given by,

$$AD(w) = \xi = (\xi_1, \ldots, \xi_N), \qquad \xi_k = \max_{0 \le j \le k}\{w_j\} - w_k \quad . \qquad 3.29$$

Note the components $w_1, \ldots, w_N$ and $(\xi_1, \ldots, \xi_N)$ of vectors w and ξ, are, time series $w_1, \ldots, w_N$ and $(\xi_1, \ldots, \xi_N)$, respectively, where the kth components of w and ξ correspond to time moment t_k. Since ξ_0 is always zero, we do not include it into drawdown time series ξ. According to Chekhlov et al. [11], $AD(w)$ and ξ are the same drawdown time series, they refer to the notation $AD(w)$ to emphasize its dependence on w.

Figure 3-1 illustrates an example of the absolute drawdown ξ and a corresponding sample path w. Figure 3-1 was adopted from Chekhlov et al. [11] and has been redrawn by the Author.

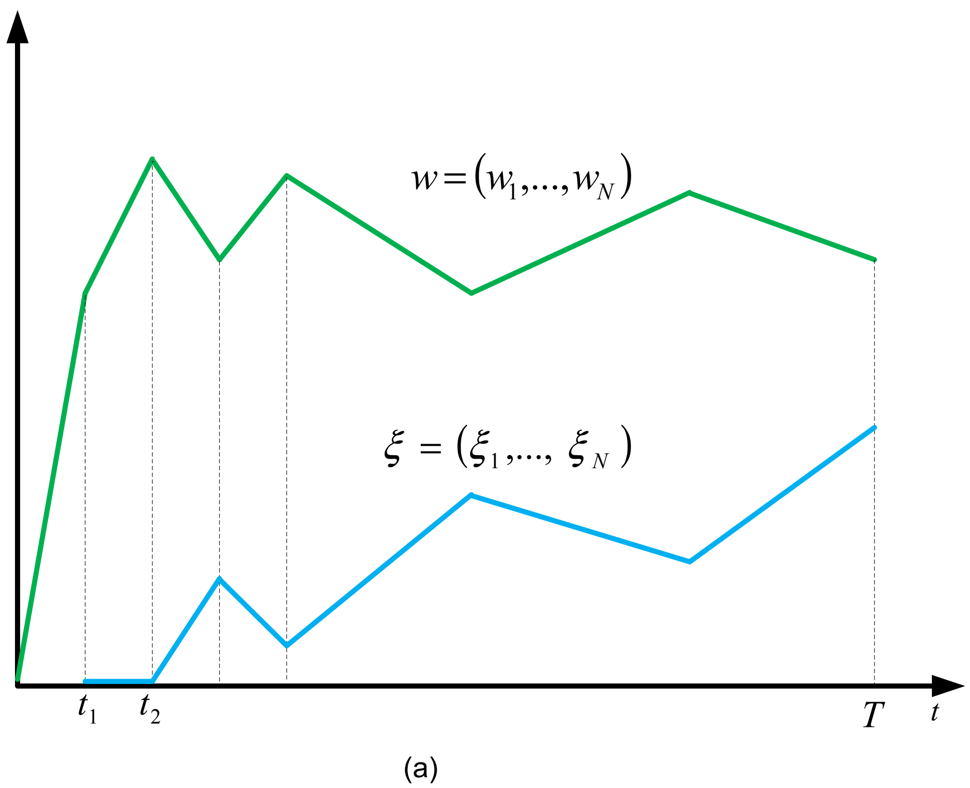


Figure 3-1: Time series of uncompounded cumulative rate of return with absolute drawdown

Proposition 3.1 [11]

Define vectorial operations as such $w + const = (w_1 + const, \ldots, w_N + const)$ and

$\lambda w = (\lambda w_1,...,\lambda w_N)$, the $AD(w)$ satisfies the following properties:

i. Nonnegativity: $AD(w) \geq 0$.

ii. Insensitivity to constant shift: $AD(w + const) = AD(w)$.

iii. Positive homogeneity: $AD(\lambda w) = \lambda AD(w), \quad \forall \lambda \geq 0$.

iv. Convexity: if $w_\lambda = \lambda w_a + (1-\lambda) w_b$ is a linear combination of any two sample paths of uncompounded cumulative rates of return, w_a and w_b, with λ ∈ [0, 1], then $AD(\lambda w) = \lambda AD(w_a) + (1-\lambda) AD(w_b)$.

3.4.2 Maximum, average and conditional drawdowns

We consider three functionals based on the notion of drawdown:

i. Maximum Drawdown (MaxDD),

ii. Average Drawdown (AvDD), and

iii. Conditional Drawdown (CDD).

The last risk functional is actually a family of performance functions depending upon parameter α. It is defined similar to CVaR [36].

Definition 3.6 [11]

For given time interval [0, T], partitioned into N subintervals $[t_{k-1}, t_k]$, $k = 1,...,N$, with $t_0 = 0$ and $t_N = T$, $AvDD$ and $MaxDD$ functionals are defined, respectively

$$AvDD(w) = \frac{1}{N}\sum_{k=1}^{N} \xi_k, \qquad 3.30$$

$$MaxDD(w) = \max_{1 \leq k \leq N} \{\xi_k\}. \qquad 3.31$$

Chekhlov et al. [11] define Conditional Value-@-Risk (CV@R) and CDD, by introducing a function $\pi_\xi(s)$ such that

$$\pi_\xi(s) = \frac{1}{N}\sum_{k=1}^{N} I_{\{\xi_{k \leq s}\}}. \qquad 3.32$$

Where $I_{\{\xi_{k \leq s}\}}$ is an indicator equal to 1, if the condition in curly brackets is true, and equal to zero, if the condition is false, i.e. with $c \in \mathbb{R}$,

$$I_{\{c \le s\}} = \begin{cases} 1, & c \le s, \\ 0, & c > s. \end{cases} \quad 3.33$$

Definition 3.7 [11]

For a given sequence of ξ_k , $k = 1,\ldots,N$, CV@R is formally defined by

$$CV@R_\alpha(\xi) = \left(\frac{\pi_\xi(\zeta(\alpha)) - \alpha}{1-\alpha}\right)\zeta(\alpha) + \frac{1}{(1-\alpha)N}\sum_{\xi_k \in \Psi}\xi_k \ . \quad 3.34$$

Where $\Psi_\alpha = \{\xi_k | \xi_k > \zeta(\alpha), k = 1,\ldots,N\}$. Let's consider the first term in the RHS of Eq. (3.34), it appears because of the inequality $\pi_\xi(\pi_\xi^{-1}(\alpha)) \ge \alpha$. If $(1-\alpha)*100\%$ of the worst drawdowns can be counted precisely, then $\pi_\xi(\pi_\xi^{-1}(\alpha)) = \alpha$ and the first term in the right-hand side of Eq. (3.34) disappears. Eq. (3.34) follows from the framework of the CVaR methodology [36]. This thus reassures the close relations between CVaR and CV@R. $CV@R_\alpha$, given by Eq. (3.34), and functional $CV@R_\alpha$, are linearly dependent. This means that if X is an arbitrary random variable then

$$CV@R_\alpha(X) = \frac{1}{1-\alpha}\left(E(X) + \alpha CV@R_\alpha(X)\right). \quad 3.35$$

Chekhlov et al. [11] use of the CV@R or CVaR is only the matter of convenience.

Definition 3.8 [11]

In a single scenario case, the CDD with tolerance level $\alpha \in [0,1]$ is the CV@R applied to the drawdown functional, $AD(w)$,

$$\Delta_\alpha(w) = CV@R_\alpha(AD(w)). \quad 3.36$$

Equivalently, interpreting ξ_k , $k = 1,\ldots,N$, to be observations of a "random variable" ξ , $\alpha -$ CDD is the of a loss function $CV@R_\alpha$.

3.4.3 Multi-scenario drawdown measure

In a multi-scenario case, CDD with tolerance level α can be interpreted as [11]:

i. The average of the worst $(1-\alpha)*100\%$ drawdowns on the drawdown surface, if the worst $(1-\alpha)*100\%$ drawdowns can be counted precisely.

ii. The linear combination of $\zeta(\alpha)$ and the average of the drawdowns strictly exceeding threshold plane $\zeta(\alpha)$, based on the fact we can precisely count of $(1-\alpha)*100\%$ drawdowns.

Definition 3.9

In a multi-scenario case, the CDD, with tolerance level $\alpha \in [0,1]$, is the multi-scenario $CV@R_\alpha$ applied to drawdown surface, $AD(w)$,

$$\Delta_\alpha(w) = CV@R_\alpha(AD(w)) \tag{3.37}$$

and drawdown measure is the mixed CDD with risk profile $\chi(\alpha)$,

$$\Delta_\alpha^+(w) = \int_0^1 \Delta_\alpha(\alpha) d\chi(\alpha) \tag{3.38}$$

where $\Delta_\alpha(w)$is given by Eq. (3.37).

In the case of discrete risk profile, drawdown measure is computed by,

$$\begin{aligned}
&\Delta_\chi^+(w) = \min_{u,y,z} \sum_{i=1}^{L} \chi_i \left(y_i + \frac{1}{(1-\alpha_i)N} \sum_{k=1}^{N} \sum_{j=1}^{K} p_j z_{ijk} \right) \\
&\text{Subject to:} \\
&z_{ijk} \geq u_{jk} - y_i, \\
&u_{jk} \geq u_{j(k-1)} - r_{jk}^p, \\
&u_{jk} \geq 0, \\
&u_{j0} = 0, \\
&z_{ijk} \geq 0, \\
&i = 1,\ldots,L, \\
&j = 1,\ldots,K, \\
&k = 1,\ldots,N.
\end{aligned} \tag{3.39}$$

3.5 Remarks

This chapter's aim was to present the theorems, properties and other propositions for the alternative risks measures to VaR and CVaR. We have presented the theorems, properties and propositions pertaining to spectral, distortion, cohorent-distortion and drawndown risk measures. This chapter serves to partially achieve Obj (I).

Chapter 4. CVaR-based portfolio optimisation models

In this chapter, we develop the portfolio optimisation model where either the objective or constraint is modelled using a CVaR as a risk measure. We shall first consider the theory of general loss distributions and its applications of CVaR in portfolio optimisation problems.

4.1 General loss distributions framework for CVaR

Measures of risk have a crucial role in optimisation under uncertainty, especially in coping with the losses that might be incurred in finance or the insurance industry. We shall show how one can get a deeper understanding of the role measures of risk play within portfolio optimisation. The basis of this subsection lies in the work of Rockafella and Uryasev [35]. Much of the concepts, definition and theorems will take its origin from Rockafella and Uryasev [35] and Alexander et al. [6].

Let's suppose loss can be represented as a multivariable function given as $z = f(x, y)$ of a decision vector representing $x \in X \subset R^n$ which may generally be a portfolio, with X expressing decision constraints, and a vector $y \in Y \subset R^m$ representing the future values of a number of variables. We take y to be random with known probability distribution. Thus z comes out as a random variable having its distribution dependent on the variable x. Any optimisation problem involving z in terms of x should then take into account expectations and the "riskiness" of x. The objective is to understand this "riskiness".

4.1.1 CVaR as a loss distribution

Suppose a random vector y is governed by a probability measure P on Y, that is independent of x. Rockafella and Uryasev [35] denote $\psi(x,\cdot)$ on R, for each x, the resulting distribution function for the loss $z = f(x, y)$, i.e.,

$$\psi(x, \zeta) = P\{y \mid f(x, y) \leq \zeta\}. \qquad 4.1$$

The following approach and assumption is that $f(x, y)$ is continuous in x and measurable in y. We then have $E\{|f(x, y)|\}$ for each $x \in X$. We denote by $\psi(x, \zeta^-)$ the left limit of $\psi(x,\cdot)$ at ζ; thus this is given by:

$$\psi(x,\zeta^-) = P\{y \mid f(x,y) < \zeta\}. \qquad 4.2$$

The difference is given by:

$$\psi(x,\zeta) - \psi(x,\zeta^-) = P\{y \mid f(x,y) = \zeta\}. \qquad 4.3$$

Eq. (4.3) is positive, so that $\psi(x,\cdot)$ has a jump at $f(x,y)$. Rockafella and Uryasev [35] called this jump, the probability "atom" at ζ. In practical applications, the confidence level $\alpha \in (0,1)$ shall be $\alpha = 0.95$ or 0.99. At this confidence level, there is a corresponding VaR, defined in the Definition 4.1 below.

Definition 4.1 [35]

The α-VaR of the loss associated with a decision x is the value:

$$\zeta_\alpha(x) = \min\{\zeta \mid \psi(x,\zeta) \geq \alpha\}. \qquad 4.4$$

The minimum in Eq. (4.4) is attained because $\psi(x,\zeta)$ is non-decreasing and right-continuous in ζ. Rockafella and Uryasev [35] consider $\psi(x,\cdot)$ as a continuous and strictly increasing, $\zeta_\alpha(x)$ is simply the unique ζ satisfying $\psi(x,\zeta) = \alpha$. If these conditions are not met, then this equation can have either no unique solution or a whole range of solutions.

Figure *4-1* has been adopted and redrawn from Rockafella and Uryasev [35] to illustrate the effects of graph $\psi(x,\cdot)$ having no solutions and a range of solutions. As in Figure 4-1 (a), with a lying in an interval of confidence levels that all yield the same VaR. The lower and upper endpoints of that interval are given respectively,

$$\alpha^-(x) = \psi\left(x,\zeta_\alpha(x)^-\right), \qquad 4.5$$

$$\alpha^+(x) = \psi\left(x,\zeta_\alpha(x)\right). \qquad 4.6$$

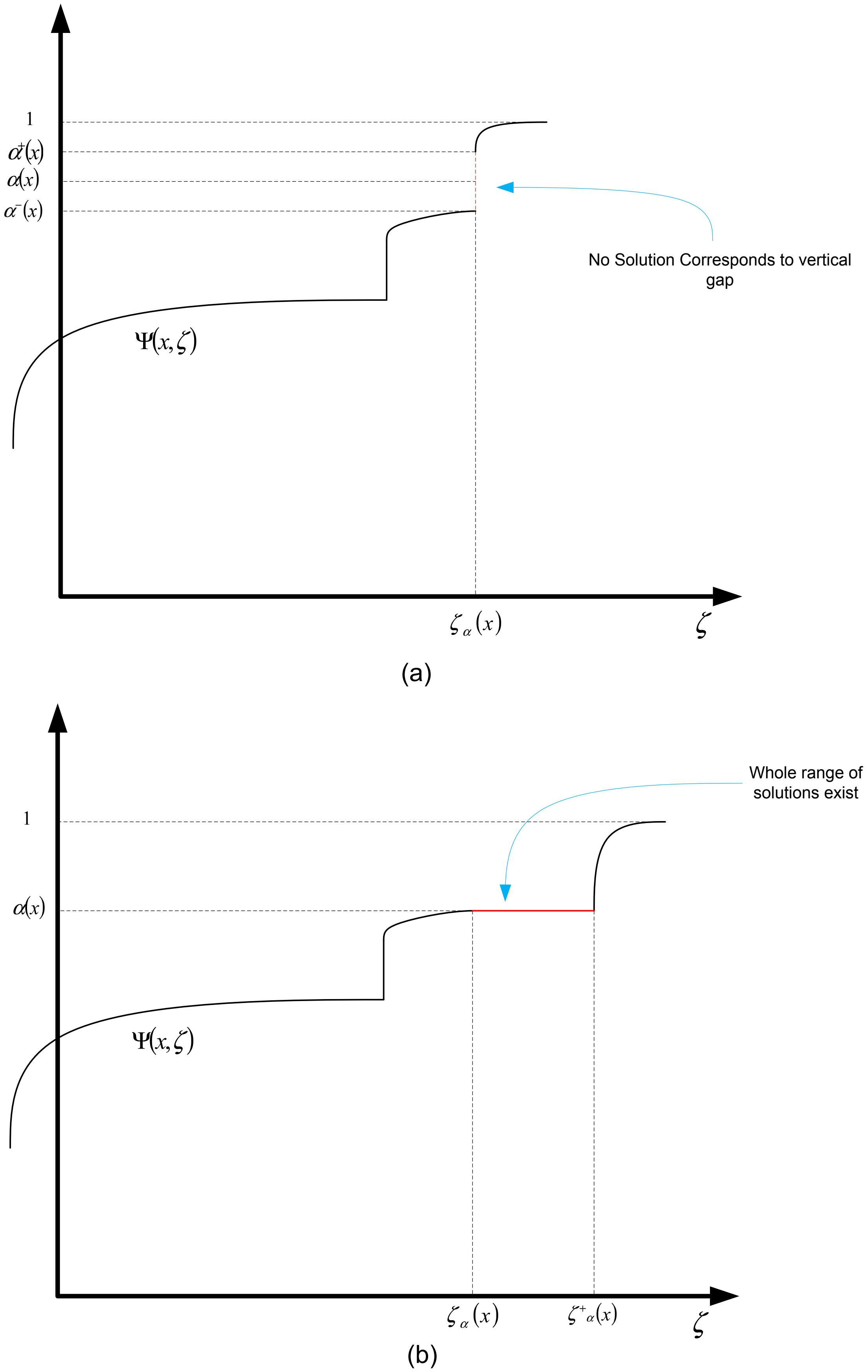


Figure 4-1: Equation with (a) no solution (b) range of solutions

With the case of a whole range of solutions, we get a constant segment of the graph, as shown in Figure 4-1(b). The solutions form an interval having $\zeta_\alpha(x)$ as its lower endpoint. The upper endpoint of the interval is the value $\zeta^+{}_\alpha(x)$.

Definition 4.2 [35]
The $\alpha - VaR^+$ also called 'upper" a-VaR, of the loss associated with a decision x is the value

$$\zeta^+{}_\alpha(x) = \inf\{\zeta | \psi(x,\zeta) > \alpha\}. \qquad 4.7$$

Note that $\zeta_\alpha(x) \le \zeta^+{}_\alpha(x)$ always holds. The values are the same except when $\psi(x,\zeta)$ is constant at level α over a certain ζ -interval. That interval is either ½$[\zeta_\alpha(x), \zeta^+{}_\alpha(x))$ or $[\zeta_\alpha(x), \zeta^+{}_\alpha(x)]$, this depends on whether or not $\psi(x,\cdot)$ has a jump at $\zeta^+_\alpha(x)$. Rockafella and Uryasev [35] in Figure 4-1 illustrates the phenomena that raise challenges in the treatment of general loss distributions. Discrete distributions associated with finite sampling or scenario modelling make for strong cases for exhibiting these types of loss distributions, since then $\psi(x,\cdot)$ is a step function.

Definition 4.3 [35]
The α-CVaR of the loss associated with a decision x is the value

$$\phi_\alpha(x) = \text{mean of the } \alpha\text{-tail distribution of } z = f(x,y). \qquad 4.8$$

Where the distribution is given with distribution function $\psi_\alpha(x,\cdot)$ defined by,

$$\psi_\alpha(x,\zeta) = \begin{cases} 0, & \text{for } \zeta < \zeta_\alpha(x). \\ \dfrac{\psi(x,\zeta)-\alpha}{1-\alpha}, & \text{for } \zeta \ge \zeta_\alpha(x). \end{cases} \qquad 4.9$$

Rockafella and Uryasev [35] note that $\psi_\alpha(x,\cdot)$ is another distribution function, in comparison $\psi_\alpha(x,\cdot)$ it is non-decreasing and right-continuous, with $\psi_\alpha(x,\zeta) \to 1$, as $\zeta \to \infty$.

Definition 4.4 [35]
The a-CVaRþ also called "upper" a-CVaR of the loss associated with a decision x is the value,

$$\phi^+_\alpha(x) = E\{f(x,y) | f(x,y) > \zeta_\alpha(x)\}. \qquad 4.10$$

Whereas the a-CVaR$^-$ also known as "lower" a-CVaR of the loss is the value,

$$\phi^-_\alpha(x) = E\{f(x,y) | f(x,y) \ge \zeta_\alpha(x)\}. \qquad 4.11$$

Rockafella and Uryasev [35] show the following basic CVaR relations. If there is no probability atom at $\zeta_\alpha(x)$, one has

$$\phi_\alpha^-(x) = \phi_\alpha(x) = \phi_\alpha^+(x). \quad 4.12$$

If a probability atom does exist at $\zeta_\alpha(x)$, one has,

$$\phi_\alpha^-(x) < \phi_\alpha(x) = \phi_\alpha^+(x) \text{ when } \alpha = \psi(x, \zeta_\alpha(x)) \quad 4.13$$

or on the other hand,

$$\phi_\alpha^-(x) = \phi_\alpha(x) \text{ when } \psi(x, \zeta_\alpha(x)) = 1. \quad 4.14$$

In all the remaining cases, characterized by

$$\psi\left(x, \zeta_\alpha(x)^-\right) < \alpha < \psi(x, \zeta_\alpha(x)) < 1. \quad 4.15$$

There is one strict inequality given by Eq. (4.16) below:

$$\phi_\alpha^-(x) < \phi_\alpha(x) < \phi_\alpha^+(x). \quad 4.16$$

For details on the proofs see Rockafella and Uryasev [35].

Rockafella and Uryasev [35] consider CVaR as a weighted average and shall be presented in the proposition below.

Proposition 4.1 [35]

Let $\mu_\alpha(x)$ be the probability assigned to the loss amount $z = \zeta_\alpha(x)$ by the α-tail distribution in Definition 4.3, namely

$$\mu_\alpha(x) = \left[\psi(x, \zeta_\alpha(x)) - \alpha\right] / [1 - \alpha] \in [0,1]. \quad 4.17$$

If $\psi(x, \zeta_\alpha(x)) < 1$, so there is chance of a loss greater than $\zeta_\alpha(x)$ then

$$\phi_\alpha(x) = \mu_\alpha(x)\zeta_\alpha(x) + \left[1 - \mu_\alpha(x)\right]\phi^+{}_\alpha(x). \quad 4.18$$

Where $\mu_\alpha(x) < 1$, whereas if $\psi(x, \zeta_\alpha(x)) = 1$, so $\zeta_\alpha(x)$ is the greatest loss that can occur, then $\phi_\alpha(x) = \zeta_\alpha(x)$.

From the definition, α-CVaR dominates a-VaR: $\phi_\alpha(x) \geq \zeta_\alpha(x)$. Also note $\phi_\alpha(x) > \zeta_\alpha(x)$ unless there is no chance of a loss greater than $\zeta_\alpha(x)$.

By representing CVaR as a certain weighted average of VaR and CVaR$^+$, Eq. (4.18) poses interesting observations. Neither VaR nor CVaR$^+$ behaves well as a measure of risk for general loss distributions. The unusual feature in the definition of CVaR that leads to its power is the way that probability atoms, can be "split" if present. Such splitting is shown in Proposition 4.1. In the notation of $\alpha^+(x)$ and $\alpha^-(x)$ in Eq. (4.5) and Eq. (4.6) and the circumstances in Eq.

(4.15), where $\alpha^{-}(x)<\alpha<\alpha^{+}(x)$. There is an atom at $\zeta_{\alpha}(x)$ having total probability $\alpha^{+}(x)-\alpha^{-}(x)$ is split into two pieces with probabilities $\alpha^{+}(x)-\alpha$ and $\alpha-\alpha^{-}(x)$ respectively. In concept, only the first of these pieces is adjoined to the interval $(\zeta_{\alpha}(x),\infty)$, which itself has probability $1-\alpha^{+}(x)$. To have to achieve a probability of $\left[1-\alpha^{+}(x)\right]+\left[\alpha^{+}(x)-\alpha\right]=1-\alpha$ we would have to choose between the intervals $\left[\zeta_{\alpha}(x),\infty\right)$ and $(\zeta_{\alpha}(x),\infty)$, of which neither actually has probability $1-\alpha$.

Rockafella and Uryasev [35] presented CVaR for scenario models and this is shown in proposition below.

Proposition 4.2 [35]
Suppose the probability measure P is concentrated in finitely many points y_k of Y, for each $x \in X$ the distribution of the loss $z=f(x,y)$ is likewise concentrated in finitely many points, and $\psi(x,\cdot)$ is a step function with jumps at those points. Fixing x, let those corresponding loss points be ordered as $z_1<z_2<...<z_N$, with the probability of z_k being $p_k>0$. Let k_a be the unique index such that

$$\sum_{k=1}^{k_\alpha} p_k \geq \alpha > \sum_{k=1}^{k_\alpha-1} p_k \,. \qquad 4.19$$

The α-VaR of the loss is given then by $\zeta_{\alpha}(x)=z_{k_\alpha}$.

Whereas the α-CVaR is given by,

$$\phi_{\alpha}(x)=\frac{1}{1-\alpha}\left[\left(\sum_{k=1}^{k_\alpha} p_k-\alpha\right)z_{k_\alpha}+\sum_{k=1}^{k_\alpha-1} p_k z_k\right]. \qquad 4.20$$

Additionally, in this situation we have,

$$\mu_{\alpha}(x)=\frac{1}{1-\alpha}\left(\sum_{k=1}^{k_\alpha} p_k-\alpha\right)\leq\frac{p_{k_\alpha}}{p_{k_\alpha}+...+p_N}. \qquad 4.21$$

For details of the proof see Rockafella and Uryasev [35].

With regards to Proposition 4.2, if the highest point z_N probability $p_N>1-\alpha$, then actually $\phi_{\alpha}(x)=\zeta_{\alpha}(x)=z_N$.

4.1.2 Coherence and formulation of a minimization rule

We now show that the $\alpha - VaR$ and $\alpha - CVaR$ of the loss z associated with a choice x can be calculated simultaneously. Additionally, by solving an elementary optimisation problem of convex type in one dimension. This subsection will show the developments from Rockafella and Uryasev [35, 36]. Rockafella and Uryasev [35] introduce the special function,

$$F_{\alpha}(x,\zeta) = \zeta + \frac{1}{1-\alpha} E\{[f(x,y) - \zeta]^{+}\}, where[t]^{+} = \max\{0,t\}. \qquad 4.22$$

Theorem 4.1 shows the minimisation formula Rockafellar and Uryasev [35] developed with special assumptions on the loss distribution. This assumption is the exclusion of discreteness. In contrast, no such formula holds for $CVaR^{+}$ or $CVaR^{-}$.

Theorem 4.1 [35]

As a function of $\zeta \in \Re, F_{\alpha}(x,\zeta)$ is finite and convex, with

$$\phi_{\alpha}(x) = \min_{\zeta} F_{\alpha}(x,\zeta). \qquad 4.23$$

The following is also given by Rockafella and Uryasev [35],

$$\zeta_{\alpha}(x) = \text{lower endpoint of } \text{argmin}_{\zeta} \mathrm{F}_{\alpha}(\mathrm{x},\zeta) \ , \qquad 4.24$$

and

$$\zeta_{\alpha}^{+}(x) = \text{upper endpoint of } \text{argmin}_{\zeta} \mathrm{F}_{\alpha}(\mathrm{x},\zeta) \ . \qquad 4.25$$

Here argmin refers to the set of ζ for which the minimum is attained. In this case has to be a nonempty, closed, bounded interval. In particular, one always has

$$\zeta_{\alpha}(x) \in \text{argmin}_{\zeta} F(x,\zeta), \qquad \phi_{\alpha} = F(x,\zeta_{\alpha}(x)). \qquad 4.26$$

For details of the proof see Rockafella and Uryasev [35].

Rockafella and Uryasev [35] show the following logic for understanding the convexity of CVaR. If $f(x,y)$ is convex with respect to x, then $\phi_{\alpha}(x)$is convex with respect to x as well. $\mathrm{F}_{\alpha}(\mathrm{x},\zeta)$ is jointly convex in (x,ζ). Likewise, $f(x,y)$ is sublinear with respect to x, then $\phi_{\alpha}(x)$ is sublinear with respect to x. Then too, $\mathrm{F}_{\alpha}(\mathrm{x},\zeta)$ is jointly sublinear in (x,ζ).

For details of the proof see Rockafella and Uryasev [35].

4.1.3 CVaR in optimisation

The optimisation of portfolios under uncertainty can either present CVaR as the objective function or a constraint and in some cases even both. The main advantage of CVaR over VaR is the preservation of convexity. In numerical applications, the joint convexity of $F_\alpha(x,\zeta)$ with respect to both x and ζ, is even more valuable than the convexity of $\phi_\alpha(x)$ in x. Rockafellar and Uryasev [35] show that the minimization of $\phi_\alpha(x)$ over $x \in \mathrm{X}$, which can be adopted as a basic prototype in the management of risk when measured by $\alpha - CVaR$. They also state that a more tractable problem is the minimisation of $F_\alpha(x,\zeta)$ in both x and ζ. We shall now present some findings noted from Rockafella and Uryasev [35, 36].

Rockafella and Uryasev [35, 36] have obtained a shortcut for the optimisation problem and the optimisation shortcut shall be presented in Theorem 4.2 below. It is important to note that the minimization of CVaR does not have to proceed numerically through repeated calculations of $\phi_\alpha(x)$for various decisions x [36]. This becomes one of the main advantages of working with CVaR as opposed to VaR. Additionally, VaR is ill-posed and has no noted shortcut in the optimisation problems.

Theorem 4.2 [35, 36]

Minimizing $\phi_\alpha(x)$ with respect to $x \in \mathrm{X}$, is equivalent to minimizing $F_\alpha(x,\zeta)$over all $(x,\zeta) \in X \times \mathbb{R}$, in the sense that

$$\min_{x \in X} \phi_\alpha(x) = \min_{(x,\zeta) \in X \text{ x } \Re} F_\alpha(x,\zeta) \qquad 4.27$$

where

$$(x^*,\zeta^*) \in \underset{(x,\zeta) \in X \text{ x } \Re}{\arg\min} F_\alpha(x,\zeta) \Leftrightarrow x^* \in \underset{x \in X}{\arg\min}\, \phi_\alpha(x),\ \zeta^* \in \underset{\zeta \in \Re}{\arg\min} F_\alpha(x^*,\zeta). \qquad 4.28$$

Rockafella and Uryasev [35, 36] show that the logic for the CVaR calculation as a by-product can be understood as follows:

If (x^*,ζ^*) minimizes F_α over $X \times \mathbb{R}$, then not only does x^* minimize ϕ_α over X, but also

$$\phi_\alpha(x^*) = F_\alpha(x^*,\zeta^*), \qquad \zeta_\alpha(x^*) \le \zeta^* \le \zeta_\alpha^+(x^*). \qquad 4.29$$

Where actually $\zeta_\alpha(x^*) = \zeta^*$ if $\arg\min_\zeta F_\alpha(x^*,\zeta)$ reduces to a single point.

It is interesting to note that $\arg\min_{\zeta} F_{\alpha}(x^*,\zeta)$ does not consist of just a single point, is possible to have $\zeta_{\alpha}(x^*) \le \zeta^*$ in Eq. (4.29). Then the joint minimization in Theorem 4.2, in producing (x^*,ζ^*), although it yields the a-CVaR associated with x^*, does not immediately yield the a-VaR associated with x^*. Further, $\arg\min_{\zeta} F_{\alpha}(x^*,\zeta)$ is the interval between two consecutive points z_k in the discrete distribution of losses. Therefore, $\zeta_{\alpha}(x^*)$ can nonetheless easily be obtained from the joint minimization: It is simply the highest $z_k \le \zeta^*$.

In Theorem 4.3, Rockafella and Uryasev [35] show that the minimization of $\phi_{\alpha}(x)$ with respect to $x \in X$ is not the only way that CVaR can be utilized. Rockafella and Uryasev [35] show in Theorem 4.3, how to "risk shape" an optimisation model. When X and g are convex and $f(x,y)$ is convex in x, we know that the optimisation problems in Theorems 4.2 and 4.3 are ones of convex programming. Due to this computation becomes very favourable. Of course, a combination of the models in Theorems 4.2 and 4.3 could likewise be handled in such a manner, by taking $g(x)=\phi_{\alpha_0}(x)$ for some α_0.

Theorem 4.3 [36]

For any selection of probability thresholds α_i and loss tolerances $\omega_i, i=1,\ldots,l$ the problem

$$\text{minimize } g(x) \text{ over } x \in X \text{ satisfying } \phi_{\alpha_i}(x) \le \omega_i \text{ for } i=1,\ldots,l. \qquad 4.30$$

Where g is any objective function chosen on X, is equivalent to the problem,

$$\begin{aligned}&\min g(x) \text{ over } (x,\zeta_1,\ldots,\zeta_l) \in X \times \mathbb{R} \times \ldots \times \mathbb{R},\\ &\text{Subject to:}\\ &\mathrm{F}_{\alpha_i}(x,\zeta_i) \le \omega, \text{ for } i=1,\ldots,l.\end{aligned} \qquad 4.31$$

Indeed, $(x^*,\zeta^*{}_1,\ldots\zeta^*{}_l)$ solves the second problem (Eq. (4.31)) if and only if x^* solves the first problem (Eq. (4.30)) and the inequality $\mathrm{F}_{\alpha_i}(x,\zeta_i) \le \omega_i$, holds for $i=1,\ldots,l$. Moreover, one then has $\phi_{\alpha_i}(x^*) \le \omega_i$ for every i, and actually $\phi_{\alpha_i}(x^*) = \omega_i$ for each i such that $\mathrm{F}_{\alpha_i}(x^*,\zeta^*{}_i) \le \omega_i$.

Linear programming techniques can be used to compute answers. That is most evident when Y is a discrete probability space with elements $y_k, k = 1,...,N$, having probabilities $p_k, k = 1,...,N$. Then we have

$$\mathrm{F}_{\alpha_i}\left(x,\zeta_i\right) = \zeta_i + \frac{1}{\left(1-\alpha_i\right)} \sum_{k=1}^{N} p_k \left[f\left(x,y_k\right) - \zeta_i \right]^+ . \qquad 4.32$$

The constraint $\mathrm{F}_{\alpha_i}\left(x,\zeta_i\right) \le \omega_i$ in Theorem 4.3 can be handled by introducing additional variables η_{ik} subject to the conditions,

$$\eta_{ik} \ge 0 \ \text{ and } f\left(x,y_k\right) - \zeta_i - \eta_{ik} \le 0 \qquad 4.33$$

and requiring that,

$$\zeta_i + \frac{1}{\left(1-\alpha_i\right)} \sum_{k=1}^{N} p_k \eta_k \le \omega_i \ . \qquad 4.34$$

The minimization in the expanded problem Eq. (4.31) is converted then into the minimization of $g(x)$ over $x \in X$, the ζ_i's and all the new η_{ik}'s, with the constraints $\mathrm{F}_{\alpha_i}\left(x,\zeta_i\right) \le \omega_i$. When f is linear in x , these constraints become linear.

4.2 Formulating the minimization of a CVaR problem

In this subsection, we shall formalise the minimization of CVaR problem. A mathematical formulation for derivative portfolio will be developed. Alexander et al. [6] assume that the available instruments $\{V_1,...,V_n\}$ are derived from the underlying assets $\{S_1,...,S_d\}$ which may be correlated. Let there be random vector $S \in \mathbb{R}^d$ to denote the underlying values.

Each derivative contract typically depends on a small subset of the underlying assets, e.g., a stock option value may depend only on one risky asset price. There are various types of derivative contracts on each underlying asset, e.g., vanilla calls and puts, exotic contracts such as binary options and barrier options with many new derivative contracts continuously emerging. For each type of option, there can be different contract specifications, e.g., strike prices and maturities, which give rise to many different possible instruments. In general, for a derivative portfolio optimisation problem, the total number of instruments n is far greater than the total number of underlying's d [6].

4.2.1 Formulating the convex programming problem

At any time t, the value of derivative contract V_i typically depends nonlinearly on the underlying. The exact value depends in the assumed model for the underlying assets and its associated parameters. For the given time horizon $\bar{t}$, let $f(x,S)$ denote the loss of portfolio with decision variable $x \in \mathbb{R}^n$ and random variable $S \in \mathbb{R}^n$ denote the value of the underlying risk factors at $\bar{t}$. Assume the probability density $p(S)$ is given by the random variable $S \in \mathbb{R}^n$. For a given portfolio x, the probability of the loss not exceeding a threshold α is given by the CDF:

$$\Psi(x,\alpha) \overset{Def}{=} \int_{f(x,S)\leq\alpha} p(S)dS. \qquad 4.35$$

Alexander et al. [6] noted that the probability distribution for the loss has no jumps, $\Psi(x,\alpha)$ is continuous everywhere with respect to α.

VaR associated with the portfolio x, for specified confidence level β and time horizon $\bar{t}$, is given by:

$$\alpha_\beta = \inf\left\{\alpha \in \mathbb{R} : \Psi(x,\alpha) \geq \beta\right\}. \qquad 4.36$$

Under the assumption the $\Psi(x,\alpha)$ is everywhere continuous and there exists α such that $\Psi(x,\alpha)=\beta$. We then define $\left[f(x,S)-\alpha\right]^+$ as

$$\left[f(x,S)-\alpha\right]^+ = \begin{cases} f(x,S), & \text{if} f(x,S)-\alpha > 0. \\ 0, & Otherwise. \end{cases} \qquad 4.37$$

The risk measure CVaR as shown by Alexander et. al [6], $\phi_\beta(x)$, is,

$$\phi_\beta(x) = \inf\left(\alpha + (1-\beta)^{-1} E\left(\left[f(x,S)-\alpha\right]^+\right)\right). \qquad 4.38$$

When the loss distribution has no jumps, CVaR is the conditional expectation of the loss, given that the loss is $\alpha_\beta(x)$ or greater, and is given by,

$$\phi_\beta(x) = (1-\beta)^{-1} \int_{f(x,S)\geq\alpha_\beta(x)} f(x,S)p(S)dS. \qquad 4.39$$

Alexander et. al [6] define the augmented function,

$$F_\beta(x,\alpha) = \alpha + (1-\beta)^{-1} \int_{S\in\mathbb{R}^d} \left[f(x,S)-\alpha\right]^+ p(S)dS. \qquad 4.40$$

The function $F_\beta(x,\alpha)$ is convex and continuously differentiable with respect to α and $\phi_\beta(x)$ is convex to x given the following assumptions are made; the loss function $f(\bullet,S)$ is convex and the loss distribution is continuous. Based on Alexander et. al [6], minimizing the CVaR over any $x \in X$, where X is a subset of $\mathbb{R}^n$, is equivalent to minimizing $F_\beta(x,\alpha)$ over $(x,\alpha) \in X \times \mathbb{R}$, i.e.

$$\min_{x \in X} \phi_\beta(x) = \min_{(x,\alpha) \in X \times \mathbb{R}} F_\beta(x,\alpha). \qquad 4.41$$

In addition, X is convex, then CVaR minimisation problem as shown by Alexander et al. [6],

$$\min_{(x,\alpha) \in X \times \mathbb{R}} F_\beta(x,\alpha) \qquad 4.42$$

is convex programming problem.

4.2.2 Minimizing the portfolio risk

We follow the assumptions made by Alexander et al. [6] and assume a given time horizon $\bar{t} \geq 0$, and that the underlying asset prices of the derivative instrument are $S_{\bar{t}} \in \mathbb{R}^d$ the initial asset prices are S_0, and the function $f(x,S)$ is the loss of a portfolio from a universe of n instruments. Assume that instrument values at time $\bar{t}$ are $\left\{V_1\left(S_{\bar{t}},\bar{t}\right),\ldots,V_n\left(S_{\bar{t}},\bar{t}\right)\right\}$. For a portfolio selection problem and a given investment horizon $\bar{t} > 0$, the loss associated with the portfolio x is

$$f(x,S_t) = -x^T\left(V^t - V^0\right). \qquad 4.43$$

Where for any time t, $V^t \overset{Def}{=} \left\{V_1(S_t,t),\ldots,V_n(S_t,t)\right\}$. Note that $f(x,S)$ is a linear function of x and it can be easily shown that, for any $q > 0$,

$$\begin{aligned} \alpha_\beta(\rho \cdot x) &= \rho \cdot \alpha_\beta(x), \\ \phi_\beta(\rho \cdot x) &= \rho \cdot \phi_\beta(x). \end{aligned} \qquad 4.44$$

Let $\delta V \in \mathbb{R}^n$ denote the change in the instrument values over the time horizon $\bar{t}$, i.e., $\delta V = V^{\bar{t}} - V^0$. Then the loss, $f\left(x,S_{\bar{t}}\right)$, of the portfolio over the investment horizon $\bar{t}$ is $-(\delta V)^T x$. Thus further emphasizing the linear relationship shown by Alexander et al. [6].

Without loss of generality, let $x \in \mathbb{R}^n$ denote the ratio of the instrument holdings to the total initial investment wealth, i.e., x_i is the number of units of the ith instrument. Assume for now that the only constraints on the optimal portfolio are the budget constraint

$$\left(V^0\right)^T x = 1 \tag{4.45}$$

and the return constraint for the investment horizon $\bar{t}$

$$\left(\overline{\delta V}\right)^T x = \bar{r} \tag{4.46}$$

where $\bar{r} \geq 0$ specifies the expected return of the portfolio over the time horizon $\bar{t}$ and $\overline{\delta V} \in \mathbb{R}^n$ is the expected gain for the instruments, i.e., $\overline{\delta V} = \mathrm{E}\left[\left(\delta V\right)\right]$.

Alexander et al. [6] show that $X = \left\{x : \left(V_0\right)^T x = 1, \left(\overline{\delta V}\right)^T x = \bar{r}\right\}$ is the set of feasible portfolios. Based on this we have,

$$\begin{aligned} &\min_{(x,\alpha)} \left(\alpha + (1-\beta)^{-1} \int_{S \in \mathbb{R}^d} \left[-(\delta V)^T x - \alpha \right]^+ p(S)\, dS \right) \\ &\text{Subject to,} \\ &\left(V^0\right)^T x = 1, \\ &\left(\overline{\delta V}\right)^T x = \bar{r}. \end{aligned} \tag{4.47}$$

Alexander et al. [6] show that one major assumption is that a stochastic model for changes of the underlying asset prices of all the instruments in a portfolio are given. They also assume that there exist methods for computing the derivative values, such as Black–Scholes formulae, delta–gamma approximations, and Monte Carlo simulation. We adopt these assumptions as well.

Delta-Gamma approach

Alexander et al. [6] investigated the delta-gamma approximation of derivative values to best understand how well an optimisation problem is posed. They used the following logical approach.

Given a short time horizon $t > 0$, a delta–gamma approximation can be a sufficiently accurate approximation to the derivative value. It is often used in risk assessment in many risk management practices. The delta–gamma approximation describes the most significant component in the change of the derivative values and can thus provide insight into the nature of the solution. Additionally, we assume for now that the change, for a given horizon $\bar{t}$, in instrument values is specified by the delta–gamma approximation for instrument i,

$$\overline{\delta V} = V_i^{\bar{t}} - V_i^0 = \left(\frac{\partial V_i^0}{\partial t}\right)\delta\bar{t} + \left(\frac{\partial V_i^0}{\partial S}\right)^T (\delta S) + \frac{1}{2}(\delta S)^T \Gamma_i (\delta S). \qquad 4.48$$

The vector $(\delta S) \in \mathbb{R}^d$ denotes the change in the underlying values, $\left(\frac{\partial V_i^0}{\partial t}\right)$ denotes the initial theta sensitivity of the i-th instrument value to time,

$$\left(\frac{\partial V_i^0}{\partial S}\right) \in \mathbb{R}^d \qquad 4.49$$

denotes the initial delta sensitivity of the i-th instrument with respect to the underlying's, and $\Gamma_i \in \mathbb{R}^{d\times d}$ is the Hessian matrix denoting the initial gamma sensitivity of the i-th instrument with respect to the underlying, and δt is change in time.

Let $\left(\frac{\partial V_i^0}{\partial t}\right)$ and $\frac{\partial V_i^0}{\partial S}$ denote the initial sensitivities for all instruments in the investment universe,

$$\begin{aligned} \frac{\partial V^0}{\partial t} &= \left[\frac{\partial V_1^0}{\partial t}, \ldots, \frac{\partial V_n^0}{\partial t}\right] \in \mathbb{R}^n, \\ \frac{\partial V^0}{\partial S} &= \left[\frac{\partial V_1^0}{\partial S}, \ldots, \frac{\partial V_n^0}{\partial S}\right] \in \mathbb{R}^{n\times d}. \end{aligned} \qquad 4.50$$

Alexander et al. [6] assume that each instrument depends on a single risky asset. If a derivative value depends on more than one risk factor, we get similar results with for the cross-partial derivatives. In the case of a single risk factor, the only non-zero entries in the vector $\frac{\partial V_i^0}{\partial S}$ and matrix $\Gamma_i \in \mathbb{R}^{d\times d}$ are entries i and (i,i), respectively. Let $\Gamma_i = \left[\Gamma^{diag}{}_1, \ldots, \Gamma^{diag}{}_n\right]^T \in \mathbb{R}^{n\times d}$, where $\Gamma^{diag}{}_i$ represents the diagonal of the matrix Γ_i as a column vector. Let $(\delta S)^2$ be the vector with each entry of δS squared. If we set

$$\Lambda = \left[\left(\frac{\partial V^0}{\partial t}\right), \left(\frac{\partial V^0}{\partial S}\right), \frac{1}{2}\Gamma\right] \in \mathbb{R}^{n\times(2d+1)} \qquad 4.51$$

the loss in portfolio value is given by

$$f(x,S) = -x^T \Lambda \begin{bmatrix} \delta \bar{t} \\ \delta S \\ (\delta S)^2 \end{bmatrix}. \quad 4.52$$

4.2.3 Derivative portfolio optimisation and its ill-posedness

Alexander et al. [6] investigate the consequences of the ill-posedness of the derivative portfolio optimisation problem. They question if these difficulties be easily overcome by imposing simple constraints, e.g., bound constraints?

We shall now present the logic on how a CVaR optimisation problem Eq. (4.47) can be solved based on the work of Alexander et al [6]. Rockafellar and Uryasev [36] introduce the auxiliary function:

$$F_\beta(x,\alpha) = \alpha + (1-\beta)^{-1} \int_{S\in\mathbb{R}^d} \left(f(x,S)-\alpha\right)^+ p(S)dS \quad 4.53$$

where

$$\left[f(x,S)-\alpha\right]^+ = \begin{cases} f(x,S), & \text{if} f(x,S)-\alpha > 0. \\ 0, & \textit{Otherwise}. \end{cases} \quad 4.54$$

It can be shown from Rockafellar and Uryasev [36] that the function $F_\beta(x,\alpha)$ is convex and continuously differentiable with respect to α if the cumulative distribution function $\Psi(x,\alpha)$ is continuous. Moreover, minimizing CVaR over any $x \in X$, where X a subset of $\mathbb{R}^n$, is equivalent to minimizing $F_\beta(x,\alpha)$ over all $(x,\alpha) \in X \in \mathbb{R}$ i.e.,

$$\min_{x\in X} \phi_\beta(x) = \min_{(x,\alpha)\in X\times\mathbb{R}} F_\beta(x,\alpha) \quad 4.55$$

The function $F_\beta(x,\alpha)$ is convex with respect to (x,α) and the CVaR function $\phi_\beta(x)$ is convex with respect to x if the loss function $f(x,S)$ is convex with respect to x. If, in addition, X is a convex set, then the minimization problem $\min_{(x,\alpha)\in X\times\mathbb{R}} F_\beta(x,\alpha)$ is a convex programming problem. The convexity property is appealing since any local minimizer of a convex programming problem is a global minimizer.

4.2.4 CVaR optimisation with cost considerations

To generate a stable solution to the CVaR optimisation problem, Alexander et al. [6] consider additional criteria since the CVaR optimisation problem for portfolio of derivatives is ill-posed. A natural meaningful consideration in portfolio investment or risk management is transaction and management cost. A portfolio that incurs a relatively small management or transactional cost is the desired outcome. Alexander et al. [6] regarded the management and transactional cost as a function of the number of instruments in a portfolio. They further discovered that it is difficult to include this explicitly into an optimisation formulation since it is computationally challenging to solve the resulting mixed integer program. Instead, they considered to find a portfolio, which consists of a small number of instruments by minimizing a combination of CVaR and a practical cost function without the need to solve a mixed integer programming problem. We adopt the definitions and logic used by Alexander et al. [6] to develop an optimisation problem with cost considerations.

We shall use the assumption from Alexander et al. [6] that the cost of holding an instrument is proportional to the magnitude of the instrument holdings. This leads to a portfolio which has a minimum weighted combination of CVaR and the proportional cost as given below:

$$\min_{x\in X}\left(\phi_\beta(x)+\sum_{i=1}^{n} c_i|x_i|\right) \qquad 4.56$$

where $\phi_\beta(x)$ is as defined in Eq. (4.39). Here $c \geq 0$ is the weighted cost, representing the cost as well as the trade-off between minimizing CVaR and cost.

The weighted cost parameter $c_i \geq 0$ can be interpreted as a measure of relative desirability to exclude the i-th instrument from the optimal portfolio. If c_i is greater than some finite threshold value, and there exists a feasible portfolio with $x_i = 0$, then the optimal portfolio x^* for Eq. (4.56) is guaranteed to exclude the i-th instrument, i.e., $x^*_i = 0$. Alexander et al. [6] show the cost model as a model for management cost. This property of the cost model Eq. (4.56) is due to the fact that the objective function $\phi_\beta(x)+\sum_{i=1}^{n} c_i|x_i|$ is an exact penalty function of a constrained optimisation problem. Alexander et al. [6] argue that if one models the cost as $\sum_{i=1}^{n} c^2_i x^2_i$ for example, the resulting optimal portfolio typically has few of its instruments with a small holding ratio $|x^*_i|$. For the quadratic penalty function, the constraint $x^*_i = 0$ is only satisfied as the penalty parameter c_i tends to $+\infty$. In order to solve Eq. (4.56), Alexander et al. consider the augmented function $F_\beta(x,\alpha)+\sum_{i=1}^{n} c_i|x_i|$. It is clear that $F_\beta(x,\alpha)+\sum_{i=1}^{n} c_i|x_i|$ remains convex and continuously differentiable with respect to α since $\sum_{i=1}^{n} c_i|x_i|$ is convex and has no dependence on α. Moreover, minimizing the sum of the weighted cost and CVaR of a portfolio x in any subset X of $\mathbb{R}^n$ is equivalent to minimizing $F_\beta(x,\alpha)+\sum_{i=1}^{n} c_i|x_i|$ over $(x,\alpha)\in X\times\mathbb{R}$. This is given by:

$$\min_{x\in X}\left(\phi_\beta(x)+\sum_{i=1}^{n} c_i|x_i|\right)\equiv\min_{(x,\alpha)\in X\times\mathbb{R}}\left(F_\beta(x,\alpha)+\sum_{i=1}^{n} c_i|x_i|\right). \qquad 4.57$$

Alexander et al. [6] show that $F_\beta(x,\alpha)+\sum_{i=1}^{n} c_i|x_i|$ is convex with respect to $(x,\alpha)\in X\times\mathbb{R}$ and $\phi_\beta(x)+\sum_{i=1}^{n} c_i|x_i|$ is convex with respect to x if the loss function $f(x,S)$ is convex with respect to x. Moreover, if X is a convex set, the minimization problem

$$\min_{(x,\alpha)\in X\times\mathbb{R}}\left(F_\beta(x,\alpha)+\sum_{i=1}^{n} c_i|x_i|\right) \qquad 4.58$$

is a convex programming problem. With Eq. (4.58) approximated through a Monte Carlo simulation, and X is specified by the budget and return constraints, the bounds on the holding ratios x, the CVaR optimisation problem with a proportional cost becomes a constrained piecewise linear minimization problem [6],

$$\min_{(x,\alpha)\in X\times\mathbb{R}}\left(\alpha+\frac{1}{m(1-\beta)}\sum_{i=1}^{m}\left[-(\delta V_i)^T x-\alpha\right]+\sum_{j=1}^{n} c_j|x_j|\right). \qquad 4.59$$

To illustrate the effect of the weighted cost parameter c on the optimal portfolio obtained from the CVaR cost model Eq. (4.58), Alexander et al. [6] consider the weighted cost parameter $c_i=\omega\cdot\left|\overline{CVaR^0}\right|, 1\le i\le n$, where $\overline{CVaR^0}$ denotes the optimal CVaR from Eq. (4.59) with no cost consideration. We shall also adopt the same weighted cost parameter in our investigations and case studies in Chapter 7.

4.2.5 Efficiency for CVaR minimization

The simulation CVaR optimisation problem Eq. (4.59) is a piecewise linear minimization problem subject to linear constraints. As discussed earlier based on results from Alexander et al. [6], one way of computing a solution to Eq. (4.59) is to solve an equivalent linear programming problem:

$$\begin{aligned}
&\min_{(x,y,z,\alpha)}\left(\alpha+\frac{1}{m(1-\beta)}\sum_{i=1}^{m}y_i+\sum_{j=1}^{n}c_j z_j\right)\\
&\text{Subject to,}\\
&\left(V^0\right)^T x=1,\\
&\left(\delta V\right)^T x=\overline{r},\\
&y\geq -Bx-\alpha e_m,\\
&z-x\geq 0,\\
&z+x\geq 0,\\
&l\leq x\leq u, y\geq 0.
\end{aligned} \qquad 4.60$$

Where the m -by-n scenario loss matrix B is given by

$$B=\left[(\delta V)_1^T;(\delta V)_2^T;\ldots;(\delta V)_m^T\right]. \qquad 4.61$$

And $e_m\in\mathbb{R}^m$ is the vector of all ones. Alexander et al. [6] conclude that this linear program has $O(n+m)$ variables and $O(n+m)$ constraints, with m as the number of Monte Carlo samples and n is the number of instruments. We assume that the loss $O(n+m)$ is computed using computational methods such as analytic formulae and/or Monte Carlo simulation.

Linear programming is the simplest constrained optimisation problem; there exists, for this class of problems, the most thorough theoretic analysis and reliable software(s). Although it is known that both CPLEX (a simplex type method) and MOSEK (an interior point method) are capable of solving very large linear programming problems in a short amount of time, the efficiency of both methods depends heavily on the sparsity structures of the problem. In addition, other programs such as MATLAB and R can solve these problems with ease on today's high end computers.

As an alternative to the linear programming approach for the CVaR optimisation problem, Alexander et al. [6] investigated a computationally efficient method which directly exploits the property of the CVaR optimisation problem. The main objective is to be able to solve large scale CVaR portfolio problems such as:

$$\min_{(x,\alpha)\in X\times\mathbb{R}}\left(F_\beta(x,\alpha)+\sum_{i=1}^{n}c_i|x_i|\right) \qquad 4.62$$

through Monte Carlo simulation. The assumption that the cumulative loss distribution function is continuous, the augmented CVaR function $F_\beta(x,\alpha)$ is continuously differentiable under the assumption that the loss distribution has no jumps. The linear programming approach arises from approximating the continuously differentiable function $F_\beta(x,\alpha)$ by the piecewise linear objective function

$$\widehat{F}_\beta(x,\alpha)=\left(\alpha+\frac{1}{m(1-\beta)}\sum_{i=1}^{m}\left[-(\delta V_i)^T x-\alpha\right]\right). \qquad 4.63$$

As the number of Monte Carlo simulations increases, the piecewise linear approximation $\widehat{F}_\beta(x,\alpha)$ approaches the continuously differentiable function $F_\beta(x,\alpha)$. An alternative to the piecewise linear approximation $\widehat{F}_\beta(x,\alpha)$, Alexander et al. [6] consider a continuously differentiable piecewise quadratic approximation $\widetilde{F}_\beta(x,\alpha)$ to the continuously differentiable function $F_\beta(x,\alpha)$. Let

$$\widetilde{F}_\beta(x,\alpha)=\left(\alpha+\frac{1}{m(1-\beta)}\sum_{i=1}^{m}\rho_\varepsilon\left[-(\delta V_i)^T x-\alpha\right]\right) \qquad 4.64$$

where $\rho_\varepsilon(z)$ is a continuously differentiable piecewise quadratic function which approximates the piecewise linear function max(z,0), given a resolution parameter $\varepsilon > 0$,

$$\rho_\varepsilon(z)=\begin{cases} z, & \text{if } z\geq\varepsilon. \\ \dfrac{z^2}{4\varepsilon}+0.5z+0.25\varepsilon, & \text{if } -\varepsilon\leq z\leq\varepsilon. \\ 0, & \text{otherwise.}\end{cases} \qquad 4.65$$

Alexander et al. [6] illustrate the smoothness of $\widehat{F}_\beta(x,\alpha)$ and $\widetilde{F}_\beta(x,\alpha)$, and we shall also consider the function $g(a)=E\left([S-\alpha]^+\right)$ assuming that S is a standard normal. We have

$$\frac{1}{m}\sum_{i=1}^{m}[S_i-\alpha]^+ \qquad 4.66$$

and $\frac{1}{m}\sum_{i=1}^{m}\rho_{\varepsilon}\left[S_i-\alpha\right]$ as compared to $g(a)$. Alexander et al. [6] show that as the number of independent samples m increases, the difference between $\frac{1}{m}\sum_{i=1}^{m}\left[S_i-\alpha\right]^{+}$ and $\frac{1}{m}\sum_{i=1}^{m}\rho_{\varepsilon}\left[S_i-\alpha\right]$, becomes smaller. In addition, the function $\frac{1}{m}\sum_{i=1}^{m}\left[S_i-\alpha\right]^{+}$ appears smoother.

Using $\widetilde{F}_{\beta}(x,\alpha)$ as a continuously differentiable approximation to $F_{\beta}(x,\alpha)$, we solve the following continuous piecewise quadratic convex programming problem,

$$\begin{aligned}
&\min_{(x,\alpha)\in X\times\mathbb{R}}\left(\widetilde{F}_{\beta}(x,\alpha)+\sum_{j=1}^{n}c_j\left|x_j\right|\right)\\
&\text{Subject to,}\\
&\left(V^0\right)^T x=1,\\
&(\delta V)^T x=\overline{r},\\
&l\le x\le u.
\end{aligned} \qquad 4.67$$

4.3 Remarks

In the first half of this chapter, we covered the concept CVaR as a general loss distribution. In particular, we represented the central concept of an probability “atom” which was introdued by Rockfellar and Uryasev [35]. We introduced the definition fo “upper” and “lower” CVaR and a definition of CVaR in the context of loss associated with a decision x. Connected to these concepts and definitions a coherence rule for minimization formula was introduced with its theorems.

In the second half of this chapter, we formulate the minimization of CVaR problem. The minimisation problem is posed as a linear program which can cater CVaR in the objective function or as a constraint. We show breifly the ill-posedness of a derivative portfolio. We then extend the CVaR optimisation problem to cater for cost considerations. Although, the first half chapter is necessary for the development of the second half of the chapter, the second half of the chapter address partially Obj (II), fully Obj (III) and fully Obj (IV). Thus Chapter 4 resulting as a key foundational chapter of the thesis.

Chapter 5. Alternative risk measures based portfolio optimisation models

In this chapter, we present how the alternative risk measures can be used in constructing a portfolio optimisation problem. As in the preceding chapter, we have developed the optimisation problem to solve for CVaR based risk measure optimisation problem. In this chapter, we develop the optimisation problem for spectral risk measures, conditional drawdown risk measures and coherent-distortion risk measures.

5.1 The minimization of general spectral measures

In this subsection, we derive the general case for minimization of spectral measure M_φ. CVaR logic and generalisation that we have already established shall form the basis for deriving the minimization of general spectral measures problem.

5.1.1 Minimization with finite number of scenarios

The estimator of a spectral measure M_ϕ shall be represented by Acerbi [1]:

$$M_\phi^N(X) = -\sum_{i=1}^{N} \phi_i X_{i:N} \,. \qquad 5.1$$

Where is ϕ_i the natural discretization of ϕ and ϕ is an admissible spectrum in a discrete sense where $\phi_i \geq 0$, $\phi_i \geq \phi_{i+1}\ \phi_i \forall_i$ and $\sum \phi_i = 1$ applies. Acerbi [1] concluded that the sorting procedure of the outcomes X_i in the general case cannot be replaced by a splitting into two subsets. All the ordered statistics $X_{i:N}$ have to be distinguished from one another, this is due to the general different weights ϕ_i. We shall not consider the generalization that takes a single auxiliary variable ψ . The general solution will require N auxiliary variables ψ_i in order to separate all the ordered statistics from one another.

Let us define $\Delta\phi_i \equiv \phi_{i+1} - \phi_i$ for $i =,\ldots, N-1$ and $\Delta\phi_N \equiv -\phi_N$. Acerbi [1] introduce a function $\Gamma_\alpha(X,\vec{\psi})$ depending on a vector $\vec{\psi} = \{\psi_1, \psi_2, \ldots, \psi_N\}$ of auxiliary variables. The function is given as:

$$\Gamma_\phi^{(N)}(X,\vec{\psi}) = \sum_{j=1}^{N} \Delta\phi_j \left\{ j\psi_j - \sum_{i=1}^{N} (\psi_j - X_i)^+ \right\}. \qquad 5.2$$

The extremal properties as shown by Adam et al. [3]:

$$
\begin{aligned}
0 &= \frac{\partial \Gamma_{\phi}^{(N)}\left(X, \vec{\psi}\right)}{\partial \psi_k} \\
&= \Delta\phi_k \left[k - \sum_{i=1}^{N} \theta\left(\psi_k - X_i\right) \right] \\
&\Leftrightarrow \begin{cases} \psi_k = \psi_k^* \in \left[X_{k:N}, X_{k+1:N}\right], & \text{if } \Delta\phi_k \neq 0, \\ \psi_k = \text{whatever}, & \text{if } \Delta\phi_k = 0. \end{cases}
\end{aligned}
\qquad 5.3
$$

Adam et al. [3] derive the following by inserting the extremal condition above into the functional,

$$
\begin{aligned}
\min_{\vec{\psi}} \left\{ \Gamma_{\varphi}^{(N)}\left(X, \vec{\psi}\right) \right\} &= \sum_{j=1}^{N} \Delta\varphi_j \left[j\psi_j^* - \sum_{i=1}^{N} \left(\psi_j^* - X_i\right)^+ \right] \\
&= \sum_{j=1}^{N} \Delta\varphi_j \left[j\psi_j^* - \sum_{i=1}^{N} \left(\psi_j^* - X_{i:N}\right)^+ \right] \\
&= \sum_{j=1}^{N} \Delta\varphi_j \left[j\psi_j^* - \sum_{i=1}^{j} \left(\psi_j^* - X_{i:N}\right) \right] \\
&= \sum_{j=1}^{N} \Delta\varphi_j \left[j\psi_j^* - j\psi_j^* \sum_{i=1}^{j} X_{i:N} \right] \\
&= \sum_{i=1}^{j} X_{i:N} \sum_{j=1}^{N} \Delta\varphi_j \\
&= \sum_{j=1}^{N} \Delta\varphi_j X_{i:N} \\
&= M_{\phi}^{N}\left(X\right).
\end{aligned}
\qquad 5.4
$$

Where $\sum_{i=1}^{N} \phi_i X_{i:N} = -\phi_i$. From Eq. (5.4) we notice that ψ_N has no part in the minimization. The minimum is always achieved for ψ_N large enough as far as $\psi_N \geq ess.\sup\{X\}$, so we can always take the limit $\psi_N \to +\infty$. Redefining $\Gamma_{\phi}^{(N)}\left(X, \vec{\psi}\right)$ as a function of $\psi_1, \psi_2, \ldots, \psi_{N-1}$ only. Adam et al. [3] define as follows:

$$
\begin{aligned}
\Gamma_{\varphi}^{(N)}\left(X, \vec{\psi}\right) &= \sum_{j=1}^{N} \Delta\varphi_j \left\{ j\psi_j - \sum_{i=1}^{N} \left(\psi_j - X_i\right) + \right\} - \lim_{\psi_N \to +\infty} \varphi_N \left\{ N\psi_N - \sum_{i=1}^{N} \left(\psi_N - X_i\right)^+ \right\} \\
&= \sum_{j=1}^{N} \Delta\varphi_j \left\{ j\psi_j - \sum_{i=1}^{N} \left(\psi_j - X_i\right) + \right\} - \varphi_N \sum_{i=1}^{N} X_i
\end{aligned}
\qquad 5.5
$$

The function $\Gamma_{\phi}^{(N)}(X,\vec{\psi})$ is convex in all its parameters $(X,\vec{\psi})$,

$$\Gamma_{\varphi}^{(N)}\left(\lambda X_1+(1-\lambda)X_2,\lambda\vec{\psi}_1+(1-\lambda)\vec{\psi}_2\right)\leq\lambda\Gamma_{\varphi}^{(N)}\left(X_1,\vec{\psi}_1\right)+(1+\lambda)\Gamma_{\varphi}^{(N)}\left(X_2,\vec{\psi}_2\right). \quad 5.6$$

For all $\lambda\in[0,1]$ provided that ϕ is an admissible risk spectrum. We summarise the following: Let $M_{\phi}^{N}(X)$ be defined by Eq. (5.1)and. Then, we have the function

$$\Gamma_{\phi}^{(N)}\left(X,\vec{\psi}\right)=\sum_{j=1}^{N}\Delta\phi_j\left\{j\psi_j-\sum_{i=1}^{N}\left(\psi_j-X_i\right)\right\}-\phi_N\sum_{i=1}^{N}X_i. \quad 5.7$$

In N-1 auxiliary parameters ψ_k is a convex, piecewise linear function in all its arguments $(X,\vec{\psi})$. Its minimum value with respect to $\vec{\psi}_k$ equals $M_{\varphi}^{N}(X)$.

5.2 Minimization of the conditional drawdown (CDD) risk measures

In this subsection, we formulate a portfolio optimisation problem with drawdown risk measure and we develop optimisation techniques for CDD efficient computation. The results and derivations shall follow the work of Chekhlov et al. [11].

5.2.1 Portfolio optimisation with drawdown measure

Chekhlov et al. [11] introduced the first requirement to calculate the value of $\zeta(\alpha)$, which causes double the computational time. Chekhlov et al. acknowledge that there is an optimisation procedure that obtains the values of threshold $\zeta(\alpha)$ and CDD simultaneously. The procedure is very important where large scale optimisation with variables and multiple constraints are encountered. When a time series of drawdowns is given, computation of the α-CDD is reduced to computation of $CV@R_{\alpha}(\xi)$.

Given a time series of instrument's drawdowns $\xi=(\xi_1,...,\xi_N)$ corresponding to time moments $\{t_1,...,t_N\}$, the CDD functional is presented by $CV@R_{\alpha}(\xi)$, which computation is reduced to the following linear programming procedure,

$$\begin{aligned}&CV@R_{\alpha}(\xi)=\min_{y,z}y+\frac{1}{(1-\alpha)N}\sum_{k=1}^{N}z_k\\&\text{Subject to,}\\&z_k\geq\xi_k-y,\\&z_k\geq0,\\&k=1,...,N.\end{aligned} \quad 5.8$$

This leading to a single optimal value of y equal to $\zeta(\alpha)$ if $\pi_{\xi}(\zeta(\alpha)) > \alpha$, and to a closed interval of optimal y with the left endpoint of $\zeta(\alpha)$ if $\pi_{\xi}(\zeta(\alpha)) = \alpha$.

For a detailed proof see Chekhlov et al. [11].

For the given knapsack problem $CV@R_{\alpha}(\xi)$ is an optimal value for the objective function:

$$
\begin{aligned}
& CDD_{\alpha}(\xi) = \min_{q} \sum_{k=1}^{N} \xi_k q_k \\
& \text{Subject to,} \\
& \sum_{k=1}^{N} q_k = 1, \\
& 0 \le q_k \le \frac{1}{(1-\alpha)N}, \\
& k = 1,\ldots,N.
\end{aligned}
\qquad 5.9
$$

The value of $CV@R_{\alpha}(\xi)$ can be found in $O(n\log_2 n)$ time. Chekhlov et al. [11] observed the knapsack problem Eq. (5.9) is dual to linear programming problem Eq. (5.8). Based on optimisation duality theory, optimal values of the objective functions in Eq. (5.8) and Eq. (5.9) should coincide. Problem Eq. (5.9) can be solved by the standard greedy algorithm in $O(n\log_2 n)$ time.

Based on the concept of a risk envelope, which is a closed, convex set of probabilities containing 1, the presentation of CV@R is related to Formulation Eq. (5.9). Suppose, a sample path of instrument's rates of return $(r_1,\ldots,r_N)$ corresponding to time moments $\{t_1,\ldots,t_N\}$, is given. For this case the uncompounded cumulative instrument's rate of return at t_k is $w_k = \sum_{l=1}^{k} r_l$, and the CDD is presented in the form of $\Delta_{\alpha}(w)$.

A sample path of instrument's rates of return $(r_1,\ldots,r_N)$, the CDD functional, $\Delta_{\alpha}(w)$, is computed by the following optimisation procedure,

$$
\begin{aligned}
&\Delta_{\alpha}(w) = \min_{u,y,z} y + \frac{1}{(1-\alpha)N}\sum_{k=1}^{N} z_k \\
&\text{Subject to,} \\
&z_k \ge u_k - y, \\
&u_k \ge u_{k-1} - r_k, \\
&u_0 = 0, \\
&u_k \ge 0, \\
&z_k \ge 0, \\
&k = 1,\dots,N.
\end{aligned}
\qquad 5.10
$$

Which leads to a single optimal value of y equal to $\zeta(\alpha)$ if $\pi_{\xi}(\zeta(\alpha)) > \alpha$, and to a closed interval of optimal y with the left endpoint of $\zeta(\alpha)$ if $\pi_{\xi}(\zeta(\alpha)) = \alpha$.

Given a sample path of instrument's rates of return $(r_1,\dots,r_N)$ the CDD functional, $\Delta_{\alpha}(w)$, is computed with optimisation procedure:

$$
\begin{aligned}
&\Delta_{\alpha}(w) = \max_{q,\eta} - \sum_{k=1}^{N} r_k \eta_k \\
&\text{Subject to,} \\
&\sum_{k=1}^{N} q_k = 1, \\
&\eta_k - \eta_{k+1} \le q_k \le \frac{1}{(1-\alpha)N}, \\
&q_k \ge 0, \eta_k \ge 0, \eta_{N+1} = 0, \\
&k = 1,\dots,N.
\end{aligned}
\qquad 5.11
$$

Given a sample path of instrument's rates of returns $\{r_k | k = 1,\dots,N\}$ and discrete risk profile $\chi_i = d\chi(\alpha_i),\ i = 1,\dots,L$ the mixed CDD, $\Delta^{+}_{\chi}(w)$ is computed by

$$
\begin{aligned}
&\Delta^{+}{}_{\chi}(w) = \min_{u,y,z} \sum_{i=1}^{L} \chi_i \left(y + \frac{1}{(1-\alpha)N} \sum_{k=1}^{N} z_{ik} \right) \\
&\text{Subject to,} \\
&z_{ik} \geq u_k - y, \\
&u_k \geq u_{k-1} - r_k, \\
&u_0 = 0, \\
&u_k \geq 0, \\
&z_{ik} \geq 0, \\
&i = 1,...,L, \\
&k = 1,...,N.
\end{aligned}
\qquad 5.12
$$

Optimal asset allocation considers the following two:

i. Generation of sample paths for the assets' rates of return.

ii. Uncompounded cumulative portfolio rate of return.

Optimal asset allocation maximizes the expected value of uncompounded cumulative portfolio rate of return at the final time moment $t_N = T$ subject to a constraint on drawdown measure. This optimisation problem is given as,

$$
\begin{aligned}
&\max_{x \in X} E_w\left(w(T,\omega,x)\right) = \sum_{j=1}^{K} p_j w_{jN}(x) \\
&\text{s.t. } \Delta_{\chi}^{+}\left(w(x)\right) \leq \gamma
\end{aligned}
\qquad 5.13
$$

where X is the set of linear "technological" constraints and $\gamma \in [0,1]$ is a proportion of the initial capital allowed to be lost.

Chekhlov et al. [11] considered an alternative approach. They considered a vector of portfolio weights to be a function of time within $[0,T]$, they assume portfolio weights $x(t_k)$ to be static for all t_k, $k = 0,...,N$. This strategy can be achieved by portfolio rebalancing at every t_k, $k = 0,...,N$. Based on the assumption made, uncompounded cumulative portfolio rate of return w is rewritten as:

$$
w_{jk}(x) = \sum_{l=1}^{k} r_{jl}^{(p)}(x) = \sum_{i=1}^{m} \sum_{l=1}^{k} r_{ij}(t_l) x_i \,. \qquad 5.14
$$

Problem in Eq. (5.8) is reduced to linear programming (LP) problem as shown below,

$$\max_{u,y,z} \sum_{j=1}^{K} p_j w_{jN}(x)$$

Subject to,

$$\sum_{i=1}^{L} \chi_i \left(y + \frac{1}{(1-\alpha)N} \sum_{k=1}^{N} z_{ik} \right) \le \gamma,$$

$$z_{ik} \ge u_{jk} - y_i,$$

$$u_{jk} \ge u_{jk-j} - r_{jk}, \qquad 5.15$$

$$u_{j0} = 0,$$

$$u_{jk} \ge 0,$$

$$z_{ijk} \ge 0,$$

$$i = 1,...,L,$$

$$j = 1,...,K,$$

$$k = 1,...,N.$$

where u_{jk} , y_i and z_{ijk} are auxiliary variables. If one considers a piece-wise function $H(x,y)$,

$$H(x,y) = \sum_{i=1}^{L} \chi_i \left(y_i + \frac{1}{(1-\alpha)N} \sum_{k=1}^{N} \sum_{j=1}^{K} p_j \left[\xi_{jk}(x) - y_i \right]^{+} \right). \qquad 5.16$$

Chekhlov et al. [11] represented the drawdown measure by:

$$\Delta_{\chi}^{+}(w(x)) = \sum_{i=1}^{L} \chi_i CV@R_{\alpha_i}(\xi(x)) = \min_{y} H(x,y). \qquad 5.17$$

Consequently, problem Eq. (5.8) is reduced to

$$\max_{x \in X} \sum_{j=1}^{K} p_j w_{jN}(x)$$
$$\text{s.t. } \min_{y} H(x,y) \le \gamma. \qquad 5.18$$

Chekhlov et la. [11] noted the most important point in their reasoning is to show that minimum in the constraint of Eq. (5.19) may be relaxed, i.e., to show that problem Eq. (5.19) is equivalent to:

$$\max_{x \in X} \sum_{j=1}^{K} p_j w_{jN}(x)$$
$$\text{s.t. } H(x,y) \le \gamma \quad . \qquad 5.19$$

Chekhlov et al. [11] proved this fact by relaxing the constraint $\min_{y} H(x,y) \le C\gamma$ in Eq. (5.19), namely, problem Eq. (5.19) is equivalently rewritten as:

$$\min_{\lambda\geq 0}\max_{x\in X}\left(\sum_{j=1}^{K}p_j w_{jN}(x)+\lambda\left(\gamma-\min_{y}H(x,y)\right)\right),$$
$$\min_{\lambda\geq 0}\max_{x\in X,y}\left(\sum_{j=1}^{K}p_j w_{jN}(x)+\lambda\left(\gamma-H(x,y)\right)\right). \quad 5.20$$

Chekhlov et al. [11] conclude that problem Eq. (5.20) is the Lagrange relaxation of Eq. (5.19). Thus, Eq. (5.19) is equivalent to Eq. (5.18).

In the cases of MaxDD(w) and AvDD(w), corresponding to the mixed CDD with risk profiles of $\chi(\alpha)=I_{\{\alpha>0\}}$ and $\chi(\alpha)=I_{\{\alpha\geq 1\}}$ LP Eq. (5.15) is simplified. The MaxDD (w) is given by:

$$\begin{aligned}
&\max_{u,y,z}\sum_{j=1}^{K}p_j w_{jN}(x)\\
&\text{Subject to,}\\
&u_{jk}\geq u_{jk-j}-r_{jk},\\
&\gamma\geq u_{jk}\geq 0,\\
&u_{j0}=0,\\
&z_{ijk}\geq 0,\\
&j=1,\ldots,K,\\
&k=1,\ldots,N.
\end{aligned} \quad 5.21$$

The AvDD(w) is given by:

$$\begin{aligned}
&\max_{u,y,z}\sum_{j=1}^{K}p_j w_{jN}(x)\\
&\text{Subject to,}\\
&u_{jk}\geq u_{jk-j}-r_{jk},\\
&\frac{1}{N}\sum_{k=1}^{N}\sum_{j=1}^{K}p_j u_{jk}\leq\gamma,\\
&u_{jk}\geq 0,\\
&u_{j0}=0,\\
&z_{ijk}\geq 0,\\
&j=1,\ldots,K,\\
&k=1,\ldots,N.
\end{aligned} \quad 5.22$$

5.2.2 Efficient frontier

Efficient frontier is a central concept in risk management methodology. Suppose for every value of γ and risk profile χ, $x_\chi^*(\gamma)$ is an optimal solution to Eq. (5.15). In this case, efficient frontier is a curve expressing dependence of optimal portfolio expected reward $\sum_{j=1}^{K} p_j w_{jN}\left(x_\chi^*(\gamma)\right)$ on portfolio risk γ.

Efficient frontier $\left(\gamma, \sum_{j=1}^{K} p_j w_{jN}\left(x_\chi^*(\gamma)\right)\right)$ is a concave curve.

Risk-adjusted return is an important characteristic for choosing an optimal portfolio on an efficient frontier that evaluates the ratio of the portfolio reward to the portfolio risk [11]:

$$\rho_\chi(\gamma) = \gamma^{-1} \sum_{j=1}^{K} p_j w_{jN}\left(x_\chi^*(\gamma)\right) \qquad 5.23$$

A fund manager is interested in such a value of $\gamma \in [0,1]$, for which the risk adjusted return $\rho_\chi(\gamma)$ is maximal. It is interpreted to be the best balance between the risk accepted and the rate of return achieved. Remembering that the efficient frontier $\left(\gamma, \sum_{j=1}^{K} p_j w_{jN}\left(x_\chi^*(\gamma)\right)\right)$ is a concave curve, $\sum_{j=1}^{K} p_j w_{jN}\left(x_\chi^*(\gamma)\right)$ is concave, hence, when this function achieves its maximum, ratio $\rho_\chi(\gamma)$ has a finite global maximum. Although $\rho_\chi(\gamma)$ is a nonlinear function with respect to γ, a problem for finding $\rho_\chi(\gamma)$ maximum and corresponding optimal γ is reduced to an LP.

The optimisation problem $\max_{\gamma \in [0,1]} \rho_\chi(\gamma)$ is reduced to LP,

$$
\begin{aligned}
&\max_{u,y,z} \sum_{j=1}^{K} p_j w_{jN}\left(\tilde{x}\right) \\
&\text{Subject to}, \\
&\sum_{i=1}^{L} \chi_i \left(\tilde{y}_i + \frac{1}{(1-\alpha)N} \sum_{k=1}^{N} p_j \tilde{z}_{ik} \right) \le 1, \\
&\tilde{z}_{ijk} \ge \tilde{u}_{jk} - \tilde{y}_i, \\
&\tilde{u}_{jk} \ge \tilde{u}_{jk-j} - r_{jk}, \\
&\tilde{u}_{j0} = 0, \\
&\tilde{u}_{jk} \ge 0, \\
&\tilde{z}_{ijk} \ge 0, \\
&i = 1,\ldots,L, \\
&j = 1,\ldots,K, \\
&k = 1,\ldots,N.
\end{aligned}
\qquad 5.24
$$

If $\tilde{x}^*$ is an optimal solution to Eq. (5.24) then $\rho_\chi\left(\gamma^*\right) = \max_{\gamma\in[0,1]} \rho_\chi\left(\gamma\right) = \sum_{j=1}^{K} p_j w_{jN}\left(\tilde{x}^*\right)$ with optimal value $\gamma^* = \frac{1}{\sum_{l=0}^{m} \tilde{x}_l^*}$ and corresponding optimal portfolio $x_l^* = \tilde{x}_l^* \gamma^*,\ l = 0,\ldots,m$.

5.3 Portfolio optimisation with coherent distortion risk measures

In this subsection, we shall consider the portfolio optimisation problem using the coherent distortion risk measure (CDRM). In this subsection, we adopt the definitions and theorems from the work Feng and Tan [15]. First, we shall recall the following special function

$$
M_g\left(x,\zeta\right) = \int_{\alpha=0}^{1} w\left(\alpha\right) F_\alpha\left(x,\zeta_\alpha\right) d\alpha. \qquad 5.25
$$

Where $w(\alpha) \ge 0$ and $\int_{\alpha=0}^{1} w(\alpha) d\alpha = 1$.

Feng and Tan [15] show that the representation of CDRM ensures the existence of $w(\alpha),\ \alpha \in [0,1]$ and defines CDRM for a given set of weights. For each α there is a corresponding auxiliary variable ζ_α. Taking partial derivatives with respect to all ζ_α for $w(\alpha),\ \alpha \in [0,1]$ and setting them equal to zeros give the extremal properties of $M_g\left(x,\zeta\right)$.

This provides more insights about the connection between a particular CDRM, $\rho_g(x)$, and its convex representation $M_g(x,\zeta)$.

Yet ζ may have infinite many entries ζ_α. Taking partial derivative with respect to all ζ_α for $w(\alpha)$, $\alpha \in [0,1]$ can be achieved by using calculus of variations. Feng and Tan [15] alleviate such difficulty by applying properties of Choquet integrals because CDRM is a subclass of DRM.

Theorem 5.1 [15]

Let $\rho_g(x)$ be a CDRM with a corresponding distortion function g. Minimizing $\rho_g(x)$ with respect to $x \in S$ is equivalent to minimizing $M_g(x,\zeta)$ over all $(x,\zeta) \in S \times \mathbb{R}^{|\zeta|}$, in the sense that

$$\min_{x\in S} \rho_g(x) = \min_{(x,\zeta)\in S\times\mathbb{R}^{|\zeta|}} M_g(x,\zeta) \tag{5.26}$$

where moreover

$$(x^*,\zeta^*) \in \arg\min M_g(x,\zeta) \Leftrightarrow x^* \in \arg\min \rho_g(x), \zeta^* \in \arg\min M_g(x^*,\zeta). \tag{5.27}$$

All results of DRM and of Choquet integrals can be applied since the CDRM is a subclass of DRM. In particular, one of the properties of Choquet integral states that if a random variable X_n has an infinite number of values and converges to X, i.e., $X_n \overset{W}{\rightarrow} X$, then $\rho_g(X_n) \overset{W}{\rightarrow} \rho(X)$ provided that $\rho_g(X)$ exists. This property implies that it is sufficient to prove the statement for the discrete random variables, and then carry over the result to the general continuous case. Consider a discrete portfolio loss random variable $l = (l_1,...,l_m)$ induced by the choice of portfolio $x \in \mathbb{R}^n$ and the random vector $y \in \mathbb{R}^m$; i.e. $l_i = l(x,y_i)$. We have the following:

$$\rho_g(x) = \sum_{i=1}^{m} w_i \phi_{\alpha i}(x). \tag{5.28}$$

Now we have the discrete analogue of Eq. (5.25) as:

$$M_g(x,) = \sum_{i=1}^{m} w_i F_{\alpha i}(x,\zeta_{\alpha i}). \tag{5.29}$$

Since $F_{\alpha i}(x,\zeta_{\alpha i})$ are all joint convex functions of x and $\zeta_{\alpha i}$, $M_g(x,\zeta)$ is a convex combination of $F_{\alpha i}(x,\zeta_{\alpha i})$, then $M_g(x,\zeta)$ is a joint convex function of x and $\zeta_{\alpha i}$. For a given portfolio x, we want to find ζ^* that minimizes $M_g(x,\zeta)$. Since $M_g(x,\zeta)$ is a convex function of $\zeta_{\alpha i}$, one can simply set the gradient of $M_g(x,\zeta)$ with respect to $\zeta_{\alpha i}$ equal to zero. Applying methodology of Feng and Tan [15] leads to:

$$
\begin{aligned}
&0 = \frac{\partial M_g(x,\zeta)}{\partial \zeta} \\
&0 = \frac{\partial}{\partial \zeta_{\alpha i}} w_i \left[\zeta_{\alpha i} + \frac{1}{1-\alpha_i} \sum_{i=1}^{m} p_i \left(l_i - \zeta_{\alpha i} \right)^+ \right] \\
&0 = w_j \left[1 - \frac{1}{1-\alpha_i} \sum_{i=1}^{m} p_i 1_{(l_i - \zeta_{\alpha i})} \right] \\
&\Leftrightarrow \begin{cases} \zeta_{\alpha i}^* \in [l_i, l_{i+1}), & \text{if } w_i \neq 0. \\ \zeta_{\alpha i}^* \text{ unconstraint}, & \text{if } w_i = 0. \end{cases}
\end{aligned}
\qquad 5.30
$$

Substituting these extremal conditions into $M_g(x,\zeta)$ Feng and Tan [15] can show that:

$$\min_{\zeta \in \mathbb{R}^m} M_g(x,\zeta) = \rho(x). \qquad 5.31$$

The minimum value of $M_g(x,\zeta)$ is precisely $\rho(x)$ and such result holds for any portfolio x. So we can replace $\rho(x)$ with $M_g(x,\zeta)$ in portfolio selection problems. Since $M_g(x,\zeta)$ is a joint convex function w.r.t (x,ζ) therefore a portfolio selection problem induces a convex programming problem if the feasible set D is convex. Based on the arguments presented above we can conclude the minimization problem for CVaR is similar to the CDRM minimization problem.

5.4 Remarks

In this chapter, we presented the optimal problems for using spectral, CDRM and drawdown risk measures. For each of the risk measures, we develop a optimal model that either can be solved by linear programming methods or genectic algorithms. This chapter addressed partially the Obj (II). Chapter 4 and Chapter 5 address Obj (II) fully.

Chapter 6. CVaR and fat tails

In this chapter, we introduce some results on CVaR and fat tails. The focus of this chapter is to introduce some popular fat tailed methodologies currently being discussed and in some cases already incorporated into risk management policies and procedures.

In the post-crisis era, there was universal agreement that financial assets [33] returns are fat-tailed and the risk managers must take extreme events into account in their risk management policies and practices. While academic research was quick to offer a vast offering of modern risk methods and analytic techniques, bringing these methods and techniques into practices needs important focus. By simply acknowledging that asset returns have higher probabilities of extreme events, has left VaR virtually useless in accurately estimating levels of risk. Skewness, auto-regression and volatility clustering must be incorporated into our modern day risk management policies and procedures.

6.1 Overview of the statistics of fat tails

When portfolio and investment managers construct practical models, they assume distributional hypothesis that captures both fat tails and asymmetry. The approach used both in academia and by practitioners is to use various different classes of distributions to capture fat tail characteristics. The most popular is the classical Student's t distribution. Some alternative examples include based on Stoyanov et al. [37]:

i. extreme value distributions,
ii. stable distributions,
iii. operator stable distributions,
iv. the class of tempered stable distributions that include stable distributions as a limiting case,
v. and the class of infinitely divisible distributions that include all previous classes except extreme value distributions.

There is no fundamental theory that can suggest a distributional model instead in remains a statistical one as noted from Rachev et al. [34]. The following is the different models based on Rachev et al. [34]:

i. Clustering of volatility- this means approximate equal price changes tend to be followed by approximate equal prices changes. (i.e. Large price changes tend to be followed by large price changes and small price changes tend to be followed by small price changes)
ii. Autoregressive behaviour-price changes depend on price changes in the past, e.g. negative price changes tend to be followed by negative price changes and vice versa.

iii. Skewness- there is an asymmetry in the upside and downside potential of price changes.
iv. Fat tails- the probability of extreme profits or losses is much larger than predicted by the normal distribution. Also the tail thickness varies for different assets. The fat tails in individual asset become relevant in portfolio return distribution as well.
v. Temporal behaviour of tail thickness- There is change through time in the probability of extreme profits and losses. The change is smaller in stable markets and larger in turbulent markets.
vi. Tail thickness varies across frequencies- high-frequency data tends to be more fat-tailed than lower-frequency data.

Except for the extreme value distributions, all these models have one key characteristic feature that they include the normal distribution as a special case. So if the data were Gaussian, the fitted distribution would be close to, or would coincide with, the normal distribution. We therefore consider these families of models to be an extension to the classical Gaussian statistical framework.

For modelling asymmetry neither the Gaussian distribution nor the classical Student's t can account for skewness. To model skewness, one can use the stable Paretian distribution, which look at the respective left and right tails. One way to capture the difference between the upside and the downside potential is by calculating expected tail loss and expected tail return. It is important to note that the degree of tail thickness varies across different asset classes.

6.2 Full distribution modelling and application of fat tail models

Risk management approaches for dealing with fat-tails includes using fat-tailed models such as Students'-t distribution, EVT, and stable Paretian distributions.

6.2.1 The Student's-t distribution

The Students t distribution is the most commonly used fat-tailed distribution as a model for asset returns by practitioners. Similarities between the normal distribution and classical Students t densities are the symmetry with a single peak. Students t exhibit densities that are more peaked around the centre and have fatter tails as compared to normal distribution. This property is illustrated on Figure 6-1. Figure 6-1 has been adopted from the work of Rachev et al. [34]. The normal distribution is a special case when the degrees of freedom (DOF) parameter approaches infinity. Rachev et al. [34] show for practical purposes, however, the Student t plot indicates a very small difference for a DOF of 30. This property makes the Student t widely accepted for various different asset returns modelling. Rachev et al. [34] note that the main advantage for using it is due to the ease of practical application with a wide range of numerical methods.

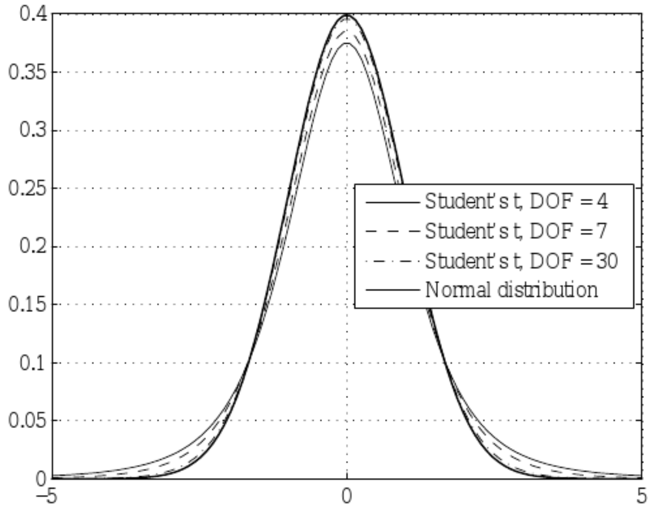


Figure 6-1: The density of Student t distribution for different DOF's [34]

It is unreasonable to accept the Gaussian distribution simply on the grounds of parsimony without any statistical analysis [37, 38]. Additionally, fixing the tail thickness to a set value for all assets makes little sense. For risk estimation simplicity, the Student’s t distribution with DOF set to five is the commonly accepted norm used by some practitioners. Some key discerning characteristics of the Gaussian model is that it is overly optimistic in times of crashes, causes significant overestimate of the risk for assets with returns that are close to being normally distributed since it fails to account for tail thickness, which varies between assets and across time.

In the Student’s t model, the fix is estimated by applying weights to the observed returns with a logarithmic or exponential decay based on a predefined parameter. This forces the relative importance of the observations in the past to be similar for all risk drivers and across all time periods. There exists an important trade-off between simplicity and precision. These models are less accurate and only work “on average" in a universe of risk drivers. Even though the approach deviates from the traditional GARCH-type framework, an implementation with the classical Student's t distribution for the residual without the deficiency of fixing the DOF parameter is available. Finally, the classical Students t model is symmetric. Rachev et al. [34] suggest there is a significant asymmetry in the data, it is advised not to be used.

6.2.2 Stable Paretian distributions

The class of stable distributions can be used as a model for asset returns because it contains the normal distribution as a special case. Only this class of distributions can approximately describe the behaviour of a stochastic system influenced by many small, regular, and independent random factors and has a distinct place among non-Gaussian full distribution models. Since price changes are driven by many random factors, it is reasonable to assume that stable distributions could represent a model for their approximate behaviour (see Stoyanov et al. [37, 38] for further details).

The tail index or index of stability is the parameter responsible for the tail behaviour in Stable Paretian distributions. Compared to DOF parameter, the index of stability is between zero and two. The closer it is to two implies the more Gaussian like the distribution is; smaller values of the index of stability imply a fatter tail. However, the Students t distribution, stable distributions allowed for skewed representatives, see Figure 6-2 below. Figure 6-2 has been adopted from the work of Rachev et al. [34].

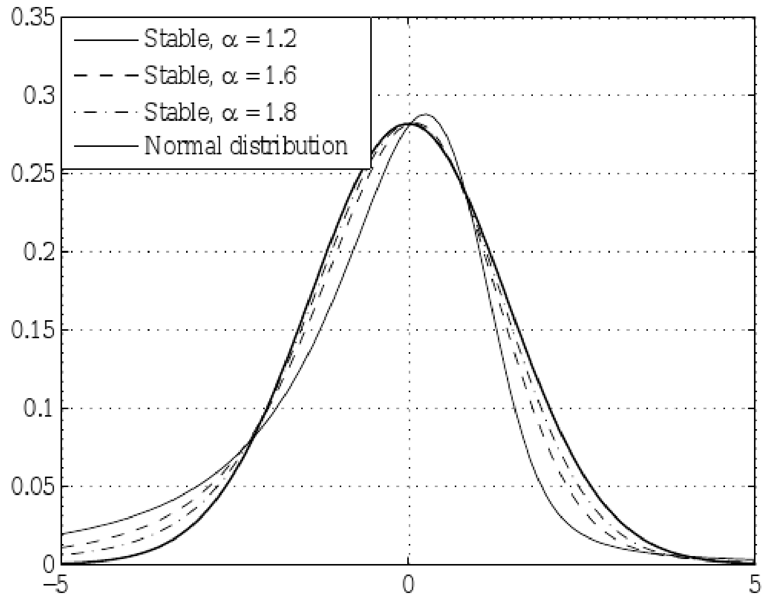


Figure 6-2: The density of the stable distribution for different index and fixed skewness [34]

Rachev et al. [34] suggest two empirical inconsistencies with the stable Paretian hypothesis. The first is that it implies infinite variance for asset returns and the second is that tail behaviour which does not change with the return frequency. These characteristics lead to very interesting developments in statistical theory. Much effort has gone into obtaining classes of distributions

which allow for aggregation across frequencies, have a finite volatility and do not deviate much from the shape of stable distributions lead to the development of tempered stable distributions. Rachev et al. [34] conclude that tempered stable distributions are derived from a special class of stable distribution through a process called tail tempering.
Tail tempering is achieved by modifying only the tails of stable distributions so that they remain thicker than the Gaussian tails but do not lead to an infinite volatility.

Rachev et al. [34] and Stoyanov et al. [37] have reached the general concept and understanding that lower frequency logarithmic returns can be represented as sums of relatively more irregular higher frequency logarithmic returns. The Gaussian behaviour of the lower frequency return can be seen as some sort of a limit behaviour, then an explanation for the change in the tail behaviour can be the convergence rate to that limit. For example, monthly returns and weekly returns can be represented as sums of daily returns, the only difference is in the number of summands. Intuitively, the convergence rate would be faster if there are more summands (higher vs lower frequencies) which are also more regular (normal vs extreme market conditions). The tail tempering technique arises from results in probability theory dealing with the problem of estimating the rates of convergence in limit theorems indicating that the shape of the distribution of the sum looks like a stable distribution at the centre but does not have as heavy tails [34, 37, 38].

6.2.3 Extreme Value Theory (EVT)- generalized Pareto distributions

EVT has applications in many other fields of science and engineering for modelling the frequency of extreme events [34]. Some extreme events include extreme temperatures, floods, winds and other natural phenomena. From a general perspective, extreme value distributions represent distributional limits for properly normalized maxima of independent random quantities with equal distributions, and therefore can be applied in finance as well [34]. In contrast to the other distribution families mentioned in the preceding subsections in this section, EVT represents a model for the tail of the distribution only. This means that in practice, one needs to combine EVT with a model for the remaining part of the distribution.
Rachev et al. [34] considered the following example to explain the idea behind EVT.
Suppose you have a sequence of returns with a given frequency. The maximum loss can be approximately defined through a limit distribution known as the generalized extreme value distribution (GED)[34, 38]. One simple way to model extreme losses is to consider the exceedances over a high threshold. EVT indicates that asymptotically, as the high threshold

increases, the exceedances can be described by the generalized Parreto distribution (GPD) [34, 38]. There are many empirical studies applying EVT directly to the return time series ignoring the clustering of volatility effect [34, 37, 38]. Rachev et al. [34] conclude and show that the two limit distributions, GEV and GPD, give rise to two approaches of EVT-based modelling, namely, the block of maxima method and the peaks-over threshold method.

Block of maxima method [34]: The idea behind the block of maxima method is twofold. The first is to take the limit behaviour described by the GEV and then divide the data into consecutive blocks of equal size and then focus on the series of maxima of returns for each block. For the second part fit the GED to the series of maxima. One method that can be applied is the method of maximum likelihood.

Peaks-over Threshold method [34]: The peaks-over-threshold (POT) method comes from the limit result leading to GPD. Compared to the block of maxima method, the parameters of GPD can be fitted using only information from the respective tail. The process is simple; choosing a value for the high threshold and fit GPD to the part of the sample which exceeds the threshold.

6.3 CVaR and fat tailed distributions

In this subsection, we present some derived results for CVaR of fat-tailed distributions from the work of Stoyanov et al [37]. CVaR is defined as the average loss provided that the loss is larger than a quantile at a given probability level,

$$CVaR_\alpha(X) = -\frac{1}{\alpha}\int_0^\alpha F^{-1}{}_X(p)dp \ . \qquad 6.1$$

Where $F^{-1}{}_X(p) = \inf\{x : P(X \le x) \ge p\}$ denotes the inverse distribution function of the random variable r_p .

Working numerically with the definition in Eq. (6.1) is difficult because the quantile function is unbounded for probabilities close to zero. Therefore, according to Stoyanov et al. [37] to calculate the integral in the definition for every distributional assumption for X. Stoyanov et al. [37] calculates the CVaR for stable distributions. The result for the symmetric case is provided in the following theorems as represented by Stoyanov et al. [37].

Theorem 6.1 [37]

If the stable distribution, $X \in S_\alpha(\sigma,0,\mu)$ with $\alpha > 1$ and $q_\varepsilon \neq 0$ is the ε-quantile of $(X-\mu)/\sigma$, then $CVaR_\varepsilon(x)$ admits the representation $CVaR_\varepsilon(x) = \sigma A_\varepsilon + \mu$, Where

$$A_\varepsilon = \frac{\alpha}{1-\alpha}\frac{|q_\varepsilon|}{\pi\varepsilon}\int_0^{\frac{\pi}{2}} g(\theta)\exp\left(-|q_\varepsilon|^{\alpha/(\alpha-1)}v(\theta)\right)d\theta. \qquad 6.2$$

Where

$$g(\theta) = \frac{\sin(\alpha-2)\theta}{\sin\alpha\theta} - \frac{\alpha\sin^2\theta}{\sin^2\alpha\theta}, \qquad 6.3$$

$$v(\theta) = \left(\frac{\cos\theta}{\sin\alpha\theta}\right)^{\alpha/\alpha-1}\frac{\cos(\alpha-1)\theta}{\cos\theta}. \qquad 6.4$$

If $q_\varepsilon = 0$, then $A_\varepsilon \frac{2\Gamma\left(\frac{\alpha-1}{\alpha}\right)}{\pi}$. A much simpler and easy to establish expression exists for the CVaR of the student's t distribution. The result is provided in the following theorem and a proof can be found in Stoyanov et al. [34, 37, 38].

Theorem 6.2 [37]

If a Student-t distribution, $X \in t(\nu,\sigma,\mu)$, with $\nu > 1$, then $CVaR_\varepsilon(x)$ admits the representation $CVaR_\varepsilon(x) = \sigma B_\varepsilon + \mu$ where,

$$B_\varepsilon = \frac{1}{\varepsilon}\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\frac{\sqrt{\nu}}{(\nu-1)\sqrt{\pi}}\left(1+\frac{q_\varepsilon^2}{\nu}\right)^{\frac{1-\nu}{2}}. \qquad 6.5$$

In which q_ε is the ε-quantile of $(X-\mu)/\sigma$. Finally, we calculate the CVaR for the GND which, as far as we know, is not available elsewhere.

Theorem 6.3 [37]

If a generalised normal distribution, $X \in GND(\kappa,\sigma,\mu)$ then for any $\varepsilon < 0.5$, $CVaR_\varepsilon(x)$ admits the representation $CVaR_\varepsilon(x) = \sigma C_\varepsilon + \mu$ where,

$$C_{\varepsilon} = \frac{1}{\varepsilon} \frac{2^{\frac{1}{\kappa}-1}}{\Gamma\left(\frac{1}{\kappa}\right)} \Gamma\left(\frac{2}{\kappa}, \frac{(-q_{\varepsilon})^{\kappa}}{2}\right). \qquad 6.6$$

In which q_{ε} is the ε-quantile of $(X-\mu)/\sigma$ and $\Gamma(s,x) = \int_{x}^{\infty} t^{s-1} e^{-t} dt$ is the upper incomplete gamma function.

6.4 CVaR and tail thickness sensitivity

In this subsection, we present the relative importance of the distribution characteristics for CVaR for Student's t distribution and stable distribution as represented in the work of Stoyanov et al. [38]. For both cases, there are expressions for CVaR which are suitable for numerical work, which can be easily applied to practical data. We apply the derived results later in Chapter 7, Case Study 4 to a set of numerical data. We use the numerical data to confirm the mathematical derivations for Student-t distributions.

6.4.1 Student's t distribution

The formula for the symmetric Student's t is [38],

$$CVaR_{\varepsilon}(X) = \begin{cases} \frac{1}{\varepsilon} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{\sqrt{\nu}}{(\nu-1)\sqrt{\pi}} \left(1 + \frac{\left(VaR_{\varepsilon}(X)\right)^{2}}{\nu}\right)^{\frac{1-\nu}{2}} - \mu, & \nu > 1. \\ \infty, & \nu = 1. \end{cases} \qquad 6.7$$

The derivatives of CVaR with respect to the distribution parameters have a structure similar to the VaR based derivatives (see Stoyanov et al. [38] for the details). Below are the given derivatives to each distribution parameter as per the work of Stoyanov et al. [38]:

$$\frac{\partial CVaR_{\varepsilon}(\upsilon,\sigma,\mu)}{\partial \upsilon} = -\sigma \frac{\partial CVaR_{\varepsilon}(\upsilon,1,0)}{\partial \nu}, \qquad 6.8$$

$$\frac{\partial CVaR_{\varepsilon}(\upsilon,\sigma,\mu)}{\partial \sigma} = CVaR_{\varepsilon}(\upsilon,1,0), \qquad 6.9$$

$$\frac{\partial CVaR_{\varepsilon}(\upsilon,\sigma,\mu)}{\partial\mu} = -1. \qquad 6.10$$

The derivative in Eq. (6.9) can be easily compared to that in Eq. (6.10). $CVaR_{\varepsilon}(\upsilon,1,0)$ is monotonic with respect to ν and ε. A numerical calculation shows that $CVaR_{0.385}(60,1,0) = 1$. Therefore, Stoyanov et al. [38] shows as a consequence,

$$\frac{\partial CVaR_{\varepsilon}(\upsilon,\sigma,\mu)}{\partial\sigma} \geq 1 \qquad 6.11$$

for $\varepsilon \leq 0.385$ and practically any ν. This result shows that the scale parameter is more efficient in changing portfolio CVaR if the tail probability is below 0.385. In comparing the derivative in Eq. (6.8), we adopt the same strategy as in the case of Student's t VaR. Stoyanov et al. [38] show Eq. (6.8) as a function of ε for three choices of ν and $\sigma = 1$.

Derivatives are higher in absolute value than the corresponding derivatives. This is due to CVaR by definition averages the quantiles in the tail which implies a higher sensitivity to tail behaviour. Even though the derivatives are generally higher, the relative importance of the tail behaviour becomes smaller than that of the mean if $\nu \geq 5$. Note that σ is at least an order of magnitude higher, we can conclude that the order of the distribution parameters by importance remains the same as in the Student's t VaR case (see [34, 37] for more details).

6.4.2 Stable distributions

The following results have been adopted from the work of Stoyanov et al. [38]. If $\alpha > 1$ and, then $VaR_{\varepsilon}(X) \neq 0$ the CVaR can be represented as $CVaR_{\varepsilon}(x) = \sigma A_{\varepsilon,\alpha,\beta} + \mu$ where the term $A_{\varepsilon,\alpha,\beta}$ does not depend on the scale and the location parameters. Concerning the term $A_{\varepsilon,\alpha,\beta}$,

$$A_{\varepsilon,\alpha,\beta} = \frac{\alpha}{1-\alpha}\frac{|VaR_{\varepsilon}(X)|}{\pi\varepsilon}\int_{-\bar{\theta}_0}^{\frac{\pi}{2}} g(\theta)\exp\left(-|VaR_{\varepsilon}(X)|^{\alpha/(\alpha-1)}\nu(\theta)\right)d\theta. \qquad 6.12$$

Where

$$g(\theta) = \frac{\sin(\alpha(\bar{\theta}_0+\theta)-2\theta)}{\sin\alpha(\bar{\theta}_0+\theta)} - \frac{\alpha\cos^2\theta}{\sin^2\alpha(\bar{\theta}_0+\theta)}, \qquad 6.13$$

$$\nu(\theta) = (\cos\alpha\bar{\theta}_0)^{\alpha/\alpha-1}\left(\frac{\cos\theta}{\sin\alpha(\bar{\theta}_0+\theta)}\right)^{\frac{\alpha}{\alpha-1}}\frac{\cos(\alpha\bar{\theta}_0+(\alpha-1)\theta)}{\cos\theta}. \qquad 6.14$$

In which $\overline{\theta}_0 = \arctan\left(\beta \tan\frac{\alpha\pi}{2}\right), \overline{\beta} = -sign(VaR_\varepsilon(X))\beta$ is the VaR of the stable distribution at tail probability ε. The derivatives of CVaR with respect to the four distribution parameters are provided below without proof from the work of Stoyanov et al. [38] and the reader is advised to refer to Stoyanov et al. [38] work for proofs:

$$\frac{\partial CVaR_\varepsilon(\alpha,\beta,\sigma,\mu)}{\partial\alpha} = -\sigma\frac{\partial CVaR_\varepsilon(\alpha,\beta,1,0)}{\partial\alpha}, \qquad 6.15$$

$$\frac{\partial CVaR_\varepsilon(\alpha,\beta,\sigma,\mu)}{\partial\beta} = -\sigma\frac{\partial CVaR_\varepsilon(\alpha,\beta,1,0)}{\partial\beta}, \qquad 6.16$$

$$\frac{\partial CVaR_\varepsilon(\alpha,\beta,\sigma,\mu)}{\partial\sigma} = \partial CVaR_\varepsilon(\alpha,\beta,1,0), \qquad 6.17$$

$$\frac{\partial CVaR_\varepsilon(\alpha,\beta,\sigma,\mu)}{\partial\mu} = -1. \qquad 6.18$$

Computing numerically the CVaR for different choices of α, β, and ε, we find out that $CVaR(\alpha,\beta,1,0) > 1$ for any $\varepsilon < 0.5$. This result implies that the scale parameter is relatively more important than the location parameter for all practical purposes. In order to compare the other derivatives, we use the information in Table 1 once again. We compute the derivatives in Eq. (6.15) and Eq. (6.16) with $\varepsilon = 0.01$ and $\sigma = 0.03$.

Chapter 7. Case Studies, simulations and results

In this chapter, we introduce case studies that will look at different problems for solving optimal portfolios. We shall consider five case studies and give the details of the problems and present our simulations and results.

The problem setting for the case studies is that we shall investigate the use of risk measures in portfolio optimisation. In Case Study 1, we shall optimise a portfolio using the CVaR, spectral and CDRM risk measures. We shall investigate and compare the solutions derived from each risk measure. We shall use the efficient frontier as a comparision mechansim when comparing the three risk measures.

In Case Study 2, we shall consider the CVaR headging problem where we hedge a target portfolio with a hedging portfolio. We shall consider a simple short call option and a long knock-out barrier option for target portfolios in two separate problems.

In Case Study 3, we shall solve a CVaR based optimisation problem where we consider cost in the objective function. We model the cost as a weighted cost parameter.

In Case Study 4, we consider performing an emperical risk analysis on the FTSE/JSE ALSI. We fit statistical distributions to the P&L distributions for various datasets of FTSE/JSE ALSI historical data. We also perform a comparative analysis using the drawdwon risk measure and CVaR.

Lastly in Case Study 5, we solve a static portfolio optimisation problem using the maximum drawdown (MaxDD), the average drawdown (AvDD), and CVaR risk measures. In this case study we solve the drawdown based optimisation problems using genetic algorithms.

7.1 Case Study 1: Comparison of risk measures in portfolio optimisation

In this case study, we demonstrate a portfolio optimisation problem were the risk measure used is CVaR. We then extend this CVaR optimisation problem to cater for the use of a spectral risk measure, with three different risk spectrums and coherent distortion risk measures (CDRM). The portfolio dataset consists of options with various characteristics such as having different volatilities, European and American styles and with various pricing methods for the derivatives. Such methods may include analytic pricing, Monte Carlo and Delta-Gamma methods. We shall use analytic pricing methods for European options and Monte Carlo methods for American options. The assets are modelled using geometric brownian motion (GBM) with no skew incorporated in the option pricing.

Dataset 1.1

Let's suppose we have a portfolio that comprises the following securities as given in Table *7-1* below. We wish to optimise this portfolio by minimizing the risk exposure for a given return of 8%. The return of 8% is an achievable outcome that is neither too aggressive nor too relaxed. It also forms a return that is above South Africa's inflation.

Table 7-1: Table of information for Dataset 1.1

Asset	Type 1	Type 2	Asset Price	Asset Strike Price	Volatility	Interest Rate	t(days)
1	Call	European	60	50	0.1	0.05	60
2	Call	European	50	50	0.13	0.05	60
3	Call	European	80	50	0.14	0.05	60
4	Put	American	50	50	0.28	0.05	60
5	Call	European	50	50	0.1	0.05	60
6	Call	European	60	50	0.16	0.05	60
7	Put	American	50	50	0.1	0.05	60
8	Call	European	50	50	0.3	0.05	60
9	Call	European	50	50	0.1	0.05	60
10	Call	European	50	50	0.1	0.05	60
11	Put	European	75	50	0.2	0.05	60
12	Put	American	100	70	0.24	0.05	60
13	Put	European	75	50	0.1	0.05	60
14	Put	European	75	50	0.23	0.05	60
15	Put	European	75	50	0.1	0.05	60
16	Put	American	75	50	0.1	0.05	60
17	Put	European	100	75	0.21	0.05	60
18	Put	European	75	50	0.1	0.05	60
19	Put	European	75	50	0.22	0.05	60
20	Put	European	100	50	0.22	0.05	60

Problem 1.1

Problem 1.1 will use the portfolio given in Dataset 1.1 and solve it using the CVaR as a risk measure. The portfolio optimisation problem is given by:

$$\min_{(x,\alpha)}\left(\alpha + m^{-1}(1-\beta)^{-1}\sum_{i=1}^{m}\left[-(\delta V)_i^T x-\alpha\right]^+\right)$$

Subject to,

$$\left(V^0\right)^T x = 1, \qquad 7.1$$

$$\left(\overline{\delta V}\right)^T x = \bar{r},$$

$$l \le x \le u.$$

The upper and lower bounds are given as 100% and -100% respectively, to account for both short and long positions.

Problem 1.2

Problem 1.2 will use the portfolio given in Dataset 1.1 and solve the following discrete case using a spectral risk measure,

$$\min_{(x,\alpha)}\left(M_\phi^N(x) = -\sum_{i=1}^{N}\varphi_i x_{i:N}\right)$$

Subject to,

$$\left(V^0\right)^T x = 1, \qquad 7.2$$

$$\left(\overline{\delta V}\right)^T x = \bar{r},$$

$$l \le x \le u.$$

The upper and lower bounds are given as 100% and -100% respectively, to account for both short and long positions. The work of Dowd and Cotter [12] considers different types of utility functions and translates them into risk spectrums. Dowd and Cotter [12] consider the following risk spectrums,

$$\phi(p) = \frac{ke^{-k(1-p)}}{1-e^{-k}}, \qquad 7.3$$

$$\phi_1(p) = \gamma(1-p)^{\gamma-1}, \text{ where } \gamma \in (0,1), \qquad 7.4$$

$$\phi_2(p) = \gamma p^{\gamma-1}, \text{ for } \gamma > 0. \qquad 7.5$$

Where $k > 0$ and γ are constants that represent risk preferences. We shall use each of the risk spectrums in solving the problem in Eq. (7.2). We wish to understand the effects of using different risk spectrums on the portfolio solutions. Based on Theorem 3.2 of Section 3, we use the spectrum generator sequence applied to each spectral risk measure in Eq. (7.3), Eq. (7.4) and Eq. (7.5),

$$\phi_i = \frac{\phi\left(i/N\right)}{\sum_{k=1}^{N}\phi\left(k/N\right)}, \qquad i = 1,\dots,N \quad . \tag{7.6}$$

This allows us to solve a discrete case for risk minimization of spectral risk measures in this problem. Many different spectrums can be used and thus may yield different optimal portfolios. In this case study, we shall explore the three risk spectrums given above.

Problem 1.3

Problem 1.3 will use the portfolio given in Dataset 1.1 and aims to solve the following using the CDRM as a risk measure,

$$\begin{aligned} &\min_{\zeta\in\mathbb{R}^m} M_g\left(x,\zeta\right) \\ &\text{Subject to,} \\ &\left(V^0\right)^T x = 1, \\ &\left(\overline{\delta V}\right)^T x = \bar{r}, \\ &l \le x \le u. \end{aligned} \tag{7.7}$$

We shall consider the following special function, the coherent-distortion fucntion, considered in Theorem 3.2 to be applied as the coherent distortion risk measure,

$$M_g\left(x,\zeta\right) = \int_{\alpha=0}^{1} w\left(\alpha\right) F_\alpha\left(x,\zeta_\alpha\right) d\alpha\ . \tag{7.8}$$

In this case study we consider the dual power function, $w(\alpha) = 1-\left(1-\alpha\right)^v$, where $v \ge 1$. Further details are given in Section 5.1.

Results

Figure *7-1* below illustrates the portfolio weighting for each asset and the position of the asset (either short or long) for using the CVaR as the risk measure. Figure *7-1* shows that there is higher weight given to short positions than long positions. European options were evaluated using the Black-Scholes analytic methods while American options were evaluted using the Longstaff-Schartz [23] Monte Carlo pricing methodology. See Appendix for a detailed example of the Longstaff-Schartz [23] Monte Carlo pricing methodology.

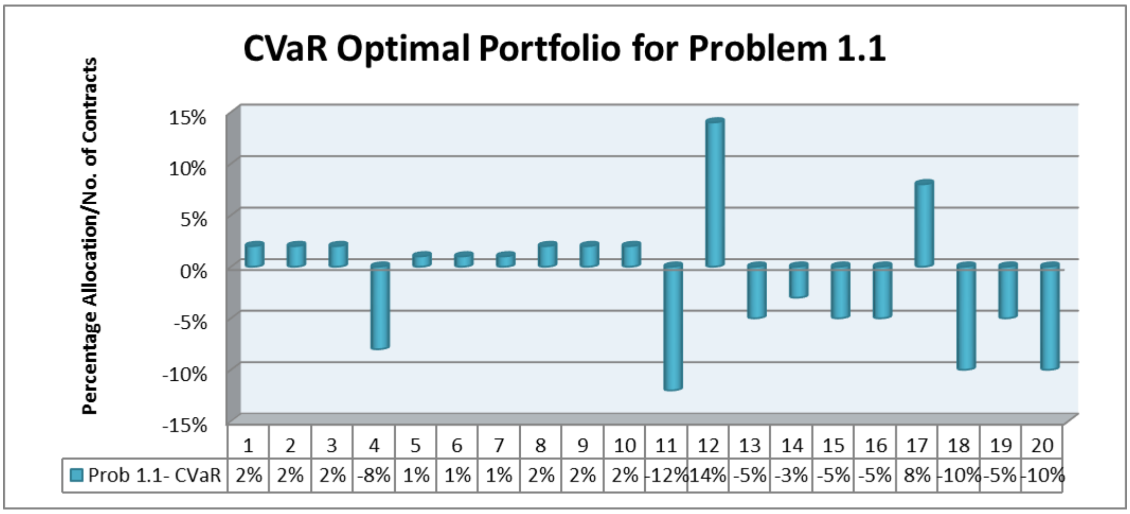


Figure 7-1: CVaR optimal portfolio for Problem 1.1

Figure *7-4* below illustrates the portfolio weighting for each asset and the position of the asset (either short or long) for using the spectral risk measure in Eq. (7.3). Figure *7-4* shows that there is higher weight given to short positions than long positions. European options were evaluated using the Black-Scholes analytic methods while American options were evaluted using the Longstaff-Schartz [23] Monte Carlo pricing methodology.

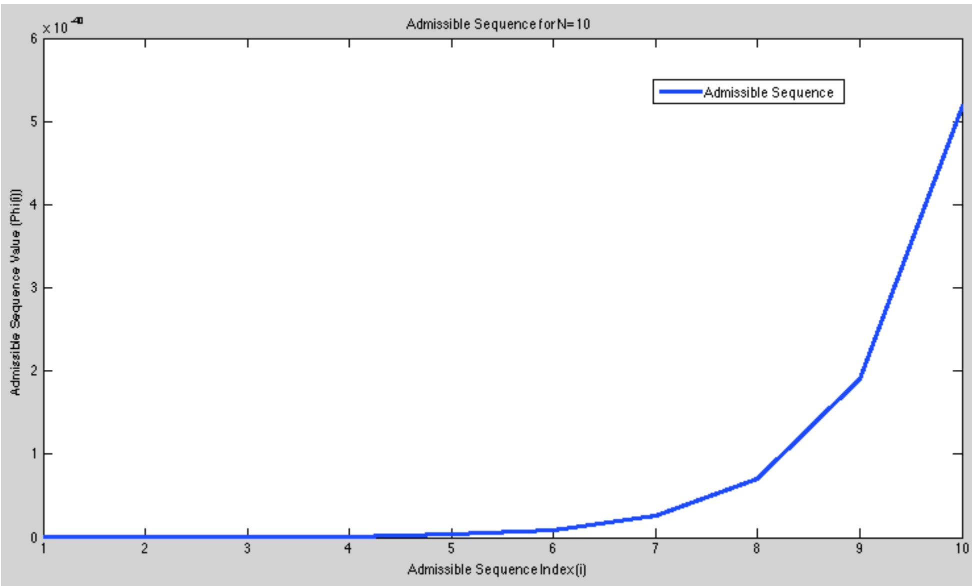


Figure 7-2: Admissible Sequence for N=10

Figure *7-2* shows the admissible sequence of the spectral measure for Eq. (7.3). Based on Theorem 3.2 of Section 3, we used the spectrum generator sequence given below,

$$\varphi_i = \frac{\varphi\left(i/N\right)}{\sum_{k=1}^{N}\varphi\left(k/N\right)}, \qquad i = 1,\ldots,N. \qquad 7.9$$

We see the decreasing values for each progression of the sequence, this in line with the work of Acerbi [1].

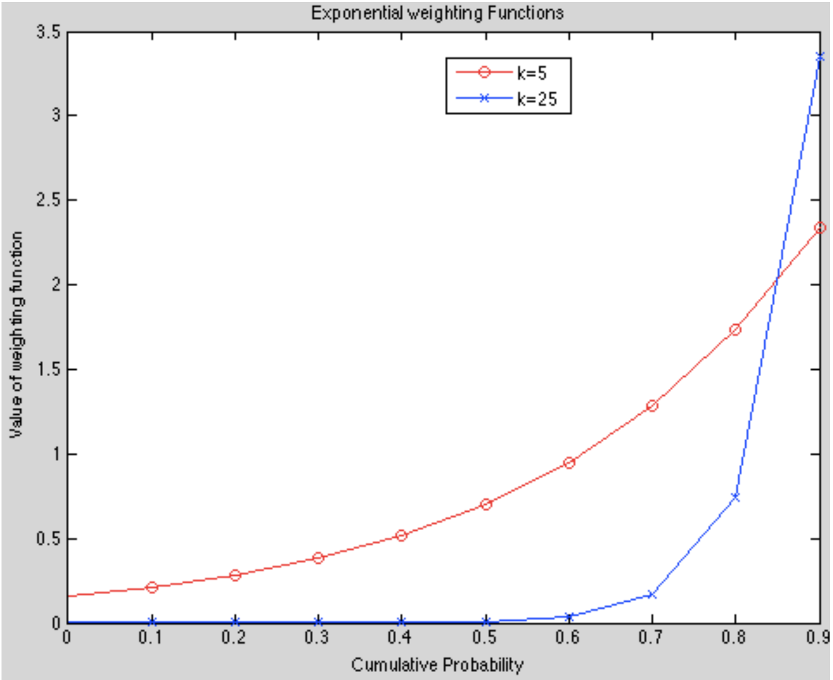


Figure 7-3: Plot for different exponential weighting function

Figure *7-3* above illustrates the plot for two different exponential weights index k, as applied to the exponential weighting function which was used as a spectral risk measure of Eq. (7.3).

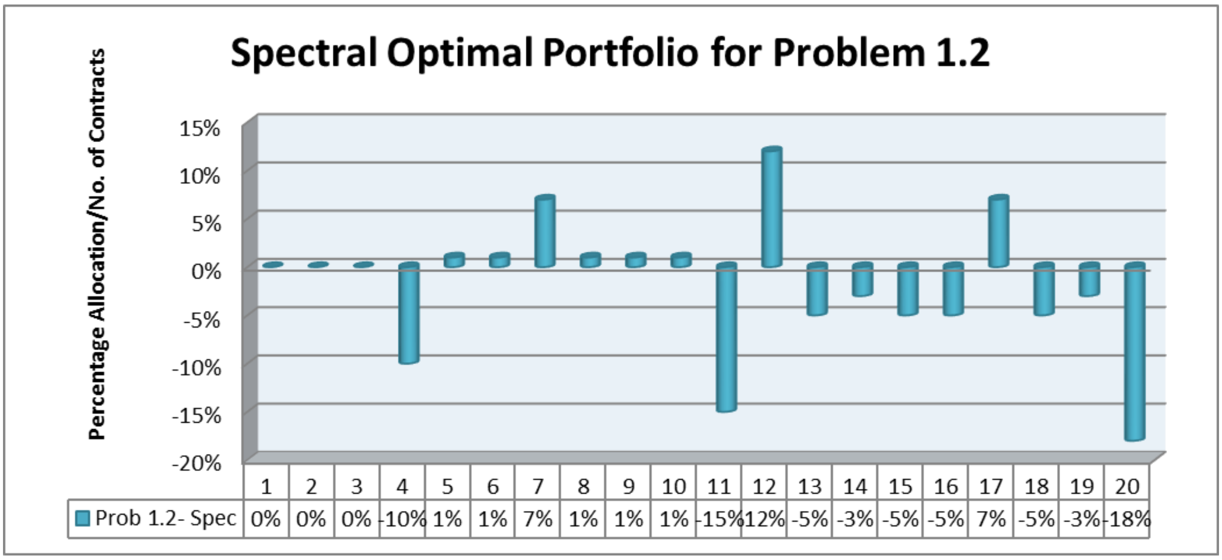


Figure 7-4: Spectral optimal portfolio for Problem 1.2 for Eq. (*7.3*)

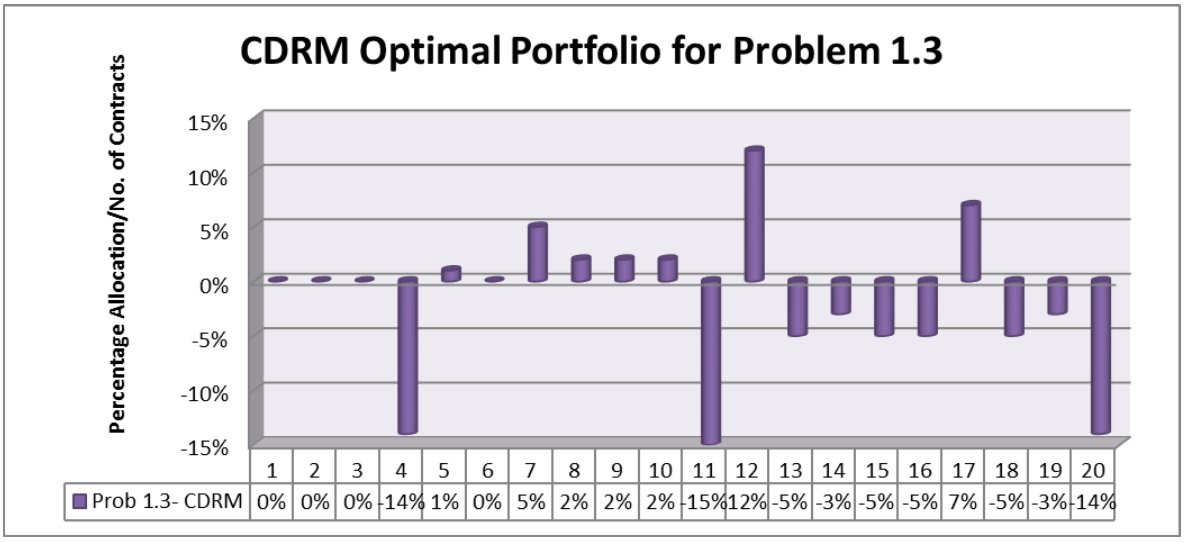


Figure 7-5: CDRM optimal portfolio for Problem 1.3

Figure *7-5* above illustrates the portfolio weighting for each asset and the position of the asset (either short or long) for using the CDRM. Figure *7-5* shows that there is higher weight given to short positions than long positions. We notice that the distribution of weights for each asset for the CVaR, Spectral and CDRM have similar patterns.

For each of the above methods we iteratively solve the problems for different return values. Based on the optimal portfolio formed we obtained the corresponding risk and thus obtained efficient frontier plots. The efficient frontier plots essentially illustrate the risk-return characteristics for each risk measure. The results from this exercise shall now be presented and discussed.

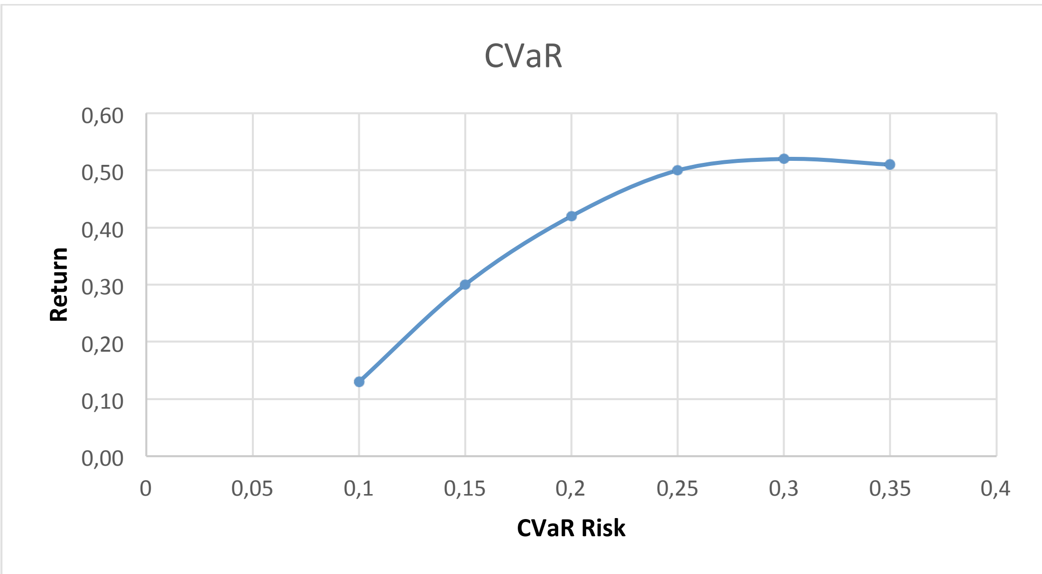


Figure 7-6: CVaR efficient frontier plot

Figure *7-6* shows the efficient frontier plot for CVaR based risk measure. The plot shows a well-defined efficient frontier. The concave looking plot ensures the commonly understood relation between risk and return for a given portfolio is preserved. However, there is slight decrease in return for increase in risk. This characteristic slightly deviates from the common characteristic. This should be noted from a traders perspective which shall also be discussed later in the case studies.

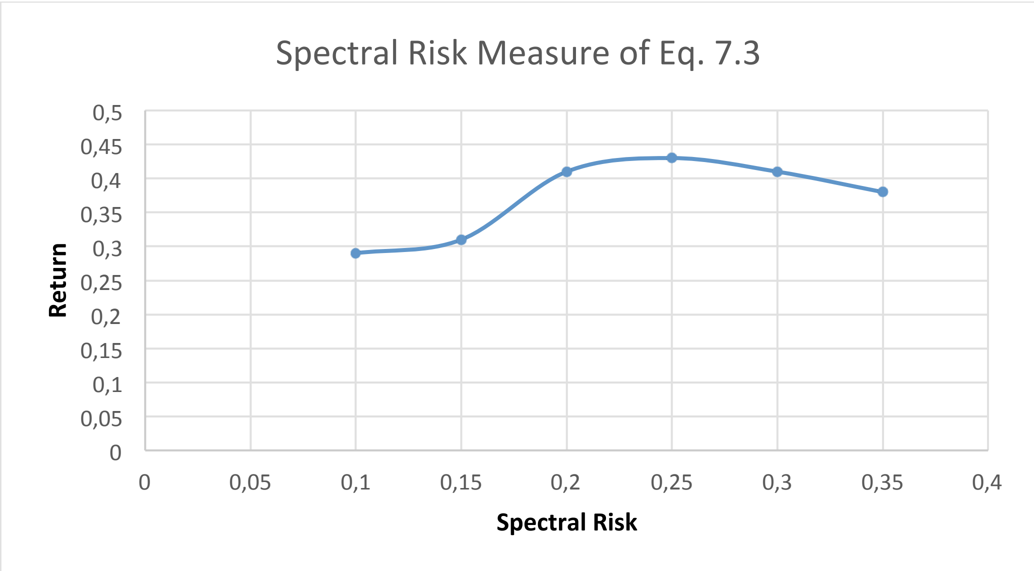


Figure 7-7: Spectral risk efficient frontier with Eq. (7.3) Spectrum

Figure *7-7* shows the efficient frontier for a spectral risk measure when using the spectrum in Eq. (7.3). The spectral risk measure efficient frontier shown in Figure *7-7*, shows some characteristics of the well understood relation between risk and return [24] but with some differences. The plot shows that for lower levers of risk the concave nature is not prominent as compared to the CVaR plot in Figure *7-6*. This poses some interesting considerations from a trading perspective, if the trader is at the lower end of the efficient frontier and wishes to increase his/her risk appetite, the expected return will not increase in the perceived fashion. The lower end of the risk plot is rather flat and then increases. This means that the trader will have to take on more risk in order to achieve his/her desired return. This issue causes some disadvantage when using the spectral risk measure as shown in Figure *7-7*. We further investigate if this characteristic exists in other spectrums. Figure *7-8* illustrates the efficient frontiers for the three spectrums used in Eq. (7.3), Eq. (7.4) and Eq. (7.5). We can see that the "flatness" at lower risk levels exist for each risk spectrum investigated. This concludes that we need to perform some further investigation(s).

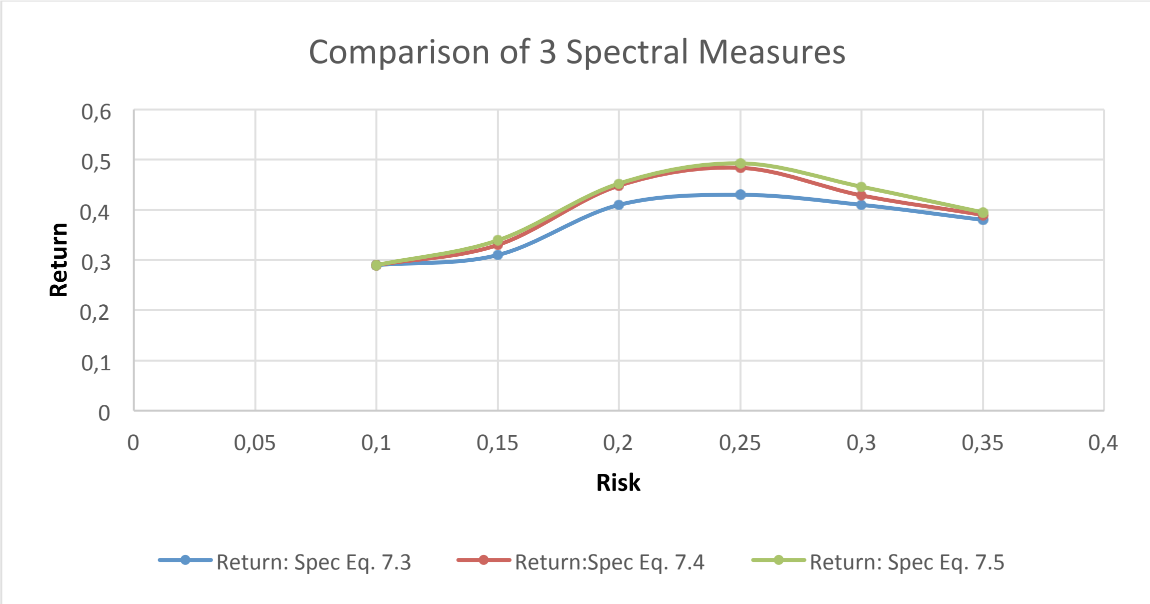


Figure 7-8: Comparison of 3 Spectral risk measure efficient frontiers

The above poses a question to understand spectral risk measures in portfolio optimisation better. The reader will question if the non-linear characteristics of derivatives or the spectral risk measure causes the lower risk to return the characteristics shown in Figure *7-7*. This side investigation shall be explored and results will be presented below in subsection 7.1.1.

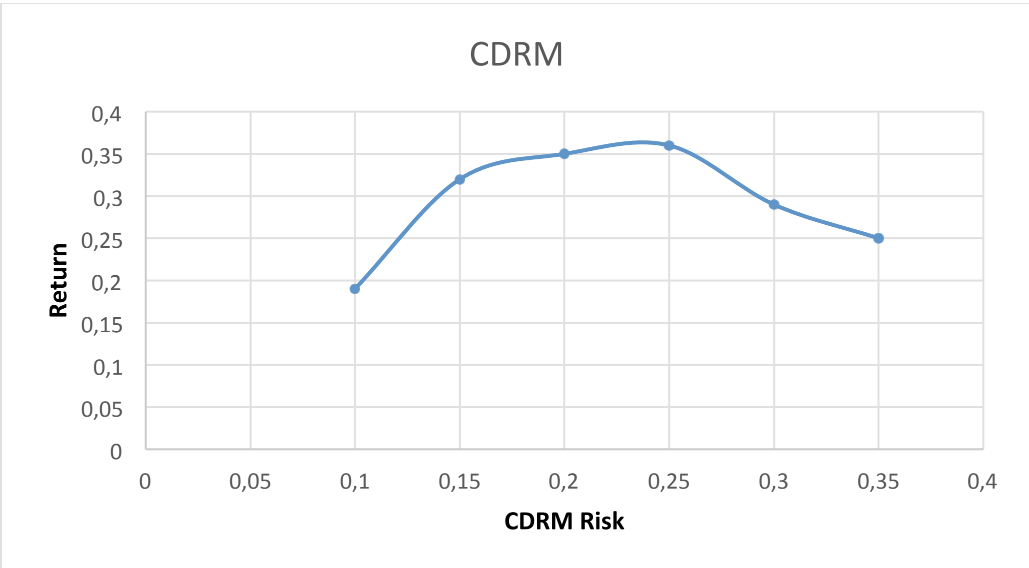


Figure 7-9: CDRM efficient frontier

Figure *7-9* shows the efficient frontier when using the coherent distortion risk measure (CDRM). The plot for the CDRM also shows similar characteristics of the generally understood relationship between risk and return [24]. In comparison with the spectral risk measure, the

CDRM shows the relatively good characteristic for lower end risk measures from a trader's perspective. However, we see that at a higher risk measure end, instead of exhibiting a flatter concave shape between risk and return, we observe that an increase in risk begins to give decrease in return. That means there is an "upper" limit of the risk-return characteristic from a trader's view. This observation means that the trader needs to trade between lower and middle ends of the risk values. This does not create that much concern as compared to the spectral risk measure unless the traders wants to have a higher risk appetite. From the efficient frontier plot in Figure *7-9*, we are satisfied with the obtained characteristics for this portfolio.

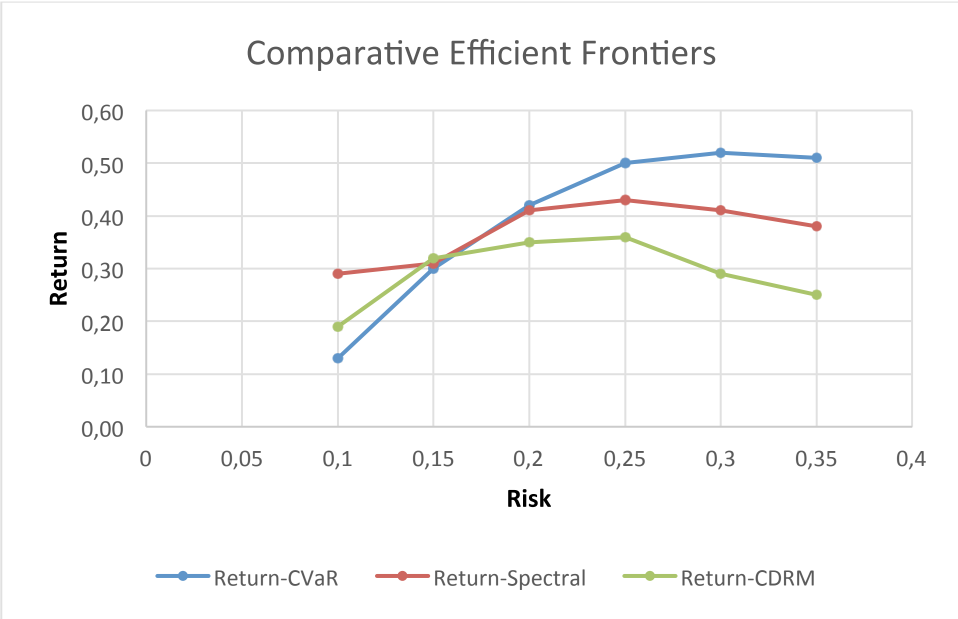


Figure 7-10: Comparison of 3 risk measures efficient frontiers

Figure *7-10* illustrates the comparison of the three risk measures efficient frontiers with the spectrum from Eq. (7.3). From Figure *7-10*, we see that for lower risk levels the CVaR and CDRM show similarities while at higher risk levels the CVaR and spectral risk measures show some similarities. From a practical point of view, both CDRM and spectral risk measures pose some difficulties when trading using the risk-return methodology.

7.1.1 Risk-Reward investigation: Comparison between spectral risk optimal portfolio and a classic Markowitz optimal portfolio

In this sub-section, we set out to understand the relation between the spectral risk optimal portfolio efficient frontier and the efficient frontier based on the Markowitz efficient frontier [24]. Acerbi and Simonetti [2] investigate a similar issue and we shall base the theoretical aspects on their work. We shall solve a linear portfolio problem where linear stocks shall be used as the underlying instruments, and apply the spectral risk measure method from the work of Acerbi [1]. We shall first briefly introduce some theoretical aspects from Acerbi and Simonetti [2].

Assume now that the profit and loss, random variable $X = X(\vec{\omega})$ depends on a set of W parameters ω_k and let $\Theta \subset \mathbb{R}^W$ be a set of acceptable weights. Based on Acerbi and Simonetti [2] the problem of minimization of a specified spectral measure $M_\phi(X(\vec{\omega}))$ with constraints $\vec{\omega} \in \Theta$ can be mapped into the equivalent minimization problem of the functional $\Gamma_\phi(X(\vec{\omega}),\psi)$. We wish to obtain the geometrical set of all optimal portfolios. Based on the work of Acerbi and Simonetti [2], the following can be setup. The $(M_\phi(X),E[X])$ optimisation problem can be naturally set up as a constrained problem where $M_\phi(X(\vec{\omega}))$ is minimized for a specified value of return $E[X]=\mu$:

$$\begin{aligned}&\min_{\omega} M_\phi(X)\\&\text{Subject to:}\\&E[X]=\mu,\\&\omega \in \Theta,\end{aligned} \qquad 7.10$$

or alternatively in the specular problem in which the expected return is maximized for a specific value $M_\phi(X(\vec{\omega}))=\rho$ of risk:

$$\begin{aligned}&\max_{\omega} E[X]\\&\text{Subject to:}\\&M_\phi(X)=\rho,\\&\omega \in \Theta.\end{aligned} \qquad 7.11$$

Now we shall consider a five stock portfolio, and solve it using the spectral risk method and compare the efficient frontier to Figure *7-7*. The Table *7-1* below shows the details of the stock considered for the linear portfolio optimisation problem using spectral risk measures.

Table 7-2: Stock portfolio details

Company Name	**Ticker**	**Time Period**	**Mean**	**Volatility**
Sasol	SOL	Jan to Dec 2015	0.03	2.5%
Old Mutual	OML	Jan to Dec 2015	0.04	2%
Anglo- American	AAL	Jan to Dec 2015	0.05	5.6%
SABMiller	SAB	Jan to Dec 2015	0.01	1.8%
FirstRand	FSR	Jan to Dec 2015	0.08	3.3%

Based on the portfolio optimisation using spectral risk measures and applying it to the linear stock shown in Table *7-2*, we obtain the results for percentage allocations. These results are illustrated in Figure *7-11* below. We illustrate results for each spectrum from Eq. (7.3), Eq. (7.4) and Eq. (7.5).

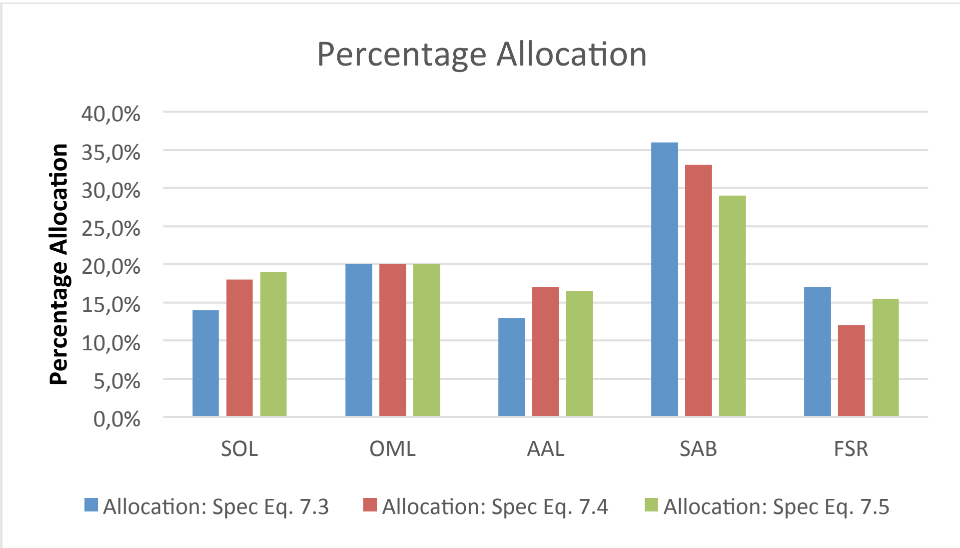


Figure 7-11: Percentage allocation for linear stock portfolio

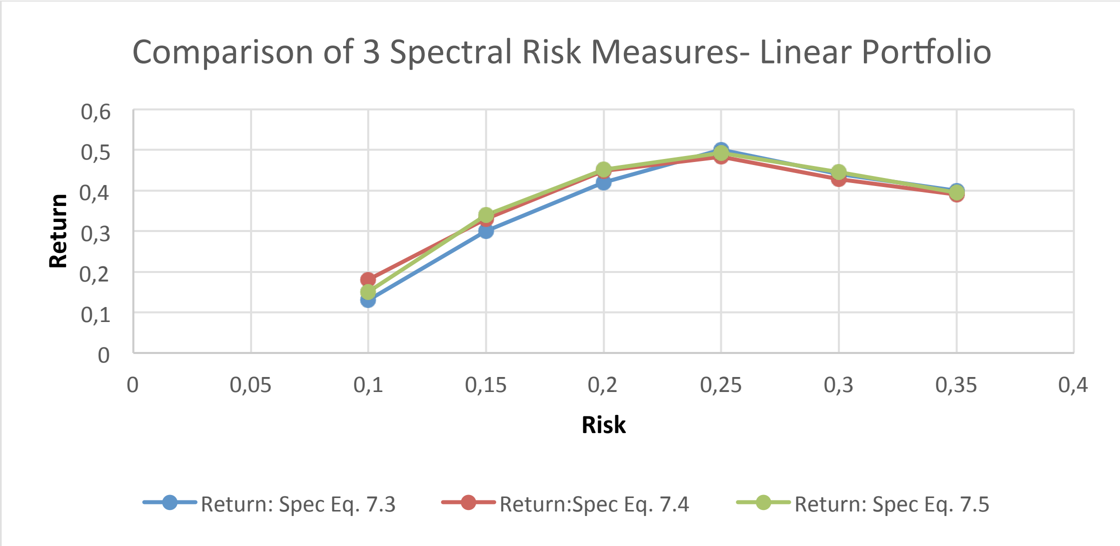


Figure 7-12: Comparison of 3 spectral risk measures for linear portfolio

Figure *7-12* illustrates the efficient frontiers for the linear stock portfolio using spectral risk measures. We plot the efficient frontiers for each risk spectrum from Eq. (7.3), Eq. (7.4) and Eq. (7.5). We pay particular attention to the lower levels of risk and compare the results from a linear portfolio, in Figure *7-12*, to that of a derivative portfolio shown in Figure *7-7* and Figure *7-8*. We can make some preliminary conclusion that the "flatness" that existed for lower levels of risk with derivative portfolios are due to the non-linear nature of derivatives. The next question to understand is: why do derivative based portfolio optimisation using spectral risk measures exhibit such characteristics on the efficient frontiers for lower level of risk? We leave this question for further investigation(s).

Remarks

In this case study we achieved solving the optimisation models using the CVaR, spectral, and CDRM risk measures. We used interior point algorithms for effectively solving the models. We presented a comparative illustration of the three risk measures from an efficient frontier perspective. We noted the salient characteristics of the concave nature of the efficient frontiers for the spectral and CDRM risk measures. We performed a deeper analysis into the flatness observed on the spectral efficient frontiers and this led to the conclusion that the flatness for lower risk values is due to the derivative behavioural characteristics and not the spectrum choices. This case study achieves the goal of Obj (II).

7.2 Case Study 2: Derivative portfolio hedging

In this case study, we hedge an option (target portfolio) by a portfolio of options. The hedging problem is to hedge a short maturity call option by trading more liquid options that are given in the datasets below. We then extend our hedging problem to hedge a long knock-out European barrier option. The datasets given below show the details for the short European call option and a long knock-out European barrier option. We will assume that there are 252 trading days in a year.

Dataset 2.1

Let's suppose we have a target portfolio that comprises of a single short call option with the following parameters as given in Table *7-3* below. We wish optimise this portfolio my minimizing the risk exposure.

Table 7-3: Detail information for target short call option

Stock	#1
Stock Price	100
Strike	100
Call/Put	Call
Long/Short	Short
Volatility	0.2
Interest Rate	0.04
Tenor	10 Days
Expected Return	0.1

Dataset 2.2

Let's suppose we have a portfolio that comprises of the following as given in Table *7-4* below for the hedging portfolio. We wish to optimise this hedging portfolio by minimising the risk exposure.

Table 7-4: Table of information for Dataset 2.2

Asset	Type 1	Type 2	Asset Price	Asset Strike Price	Volatility	Interest Rate	t(days)
1	Call	European	60	50	0.1	0.05	60
2	Call	European	50	50	0.13	0.05	60
3	Call	European	80	50	0.14	0.05	60
4	Put	American	50	50	0.28	0.05	60
5	Call	European	50	50	0.1	0.05	60
6	Call	European	60	50	0.16	0.05	60
7	Put	American	50	50	0.1	0.05	60
8	Call	European	50	50	0.3	0.05	60
9	Call	European	50	50	0.1	0.05	60
10	Call	European	50	50	0.1	0.05	60
11	Put	European	75	50	0.2	0.05	60
12	Put	American	100	70	0.24	0.05	60
13	Put	European	75	50	0.1	0.05	60
14	Put	European	75	50	0.23	0.05	60
15	Put	European	75	50	0.1	0.05	60
16	Put	American	75	50	0.1	0.05	60
17	Put	European	100	75	0.21	0.05	60
18	Put	European	75	50	0.1	0.05	60
19	Put	European	75	50	0.22	0.05	60
20	Put	European	100	50	0.22	0.05	60

Dataset 2.3

Let's suppose we have a target portfolio that comprise a single long knock-out European barrier call option with the following parameters as given in Table *7-5* below.

Table 7-5: Table of information for barrier option

Stock	#2
Stock Price	100
Strike	115
Type	KO European Barrier
Knock-out Barrier	70
Call/Put	Call
Long/Short	Long
Volatility	0.2
Interest Rate	0.04
Tenor	5 Yr. (1260 Days)
Expected Return	0.1

Dataset 2.4

Table *7-6* below shows the information for the hedging portfolio. The hedging portfolio consists of both European and American style options. The expiry dates consist of 1 month, 2 months, 3 months, 6 months and 1 year. We wish to minimise risk exposure of the Barrier option for a short time period, (upto a year) in Dataset 2.3 with the options presented below.

Table 7-6: Table of information for Dataset 2.4

Asset	Type 1	Type 2	Asset Price	Asset Strike	Volatility	Interest Rate	t(days)
1	Call	European	60	50	0.1	0.05	60
2	Call	European	50	50	0.13	0.05	30
3	Call	European	80	50	0.14	0.05	90
4	Put	American	50	50	0.28	0.05	60
5	Call	European	50	50	0.1	0.05	30
6	Call	European	60	50	0.16	0.05	90
7	Put	American	50	50	0.1	0.05	60
8	Call	European	50	50	0.3	0.05	60
9	Call	European	50	50	0.1	0.05	60
10	Call	European	50	50	0.1	0.05	252
11	Put	European	75	50	0.2	0.05	60
12	Put	American	100	70	0.24	0.05	252
13	Put	European	75	50	0.1	0.05	60
14	Put	European	75	50	0.23	0.05	90
15	Put	European	75	50	0.1	0.05	252
16	Put	American	75	50	0.1	0.05	60
17	Put	European	100	75	0.21	0.05	60
18	Put	European	75	50	0.1	0.05	126
19	Put	European	75	50	0.22	0.05	60
20	Put	European	100	50	0.22	0.05	60

Problem 2.1

In problem 2.1 we shall hedge the target portfolio given in dataset 2.1 with the portfolio given in dataset 2.2. We solve the following problem,

$$\min_{(x,\alpha)}\left(\alpha + m^{-1}(1-\beta)^{-1}\sum_{i=1}^{m}\left[\left(\Pi^0\right)_i - (\delta V)_i^T x - \alpha\right]^+\right)$$

$$\begin{aligned}&\text{Subject to:}\\&\left(V^0\right)^T x = 1,\\&\left(\overline{\delta V}\right)^T x = \bar{r},\\&l \le x \le u.\end{aligned} \qquad 7.12$$

The hedging horizon t is the maturity of the call and the loss of the existing portfolio is,

$$\Pi^0(S,t) = P_0^{\text{int}} + \max(S_t - K, 0)\,. \qquad 7.13$$

Problem 2.2

In Problem 2.2, we shall hedge the target portfolio given in Dataset 2.3 with the portfolio given in Dataset 2.4. We solve the following problem:

$$\begin{aligned} &\min_{(x,\alpha)}\left(\alpha + m^{-1}(1-\beta)^{-1}\sum_{i=1}^{m}\left[\left(\Pi^0\right)_i - (\delta V)_i^T x - \alpha\right]^+\right) \\ &\text{Subject to:} \\ &\left(V^0\right)^T x = 1, \qquad 7.14 \\ &\left(\overline{\delta V}\right)^T x = \bar{r}, \\ &l \le x \le u. \end{aligned}$$

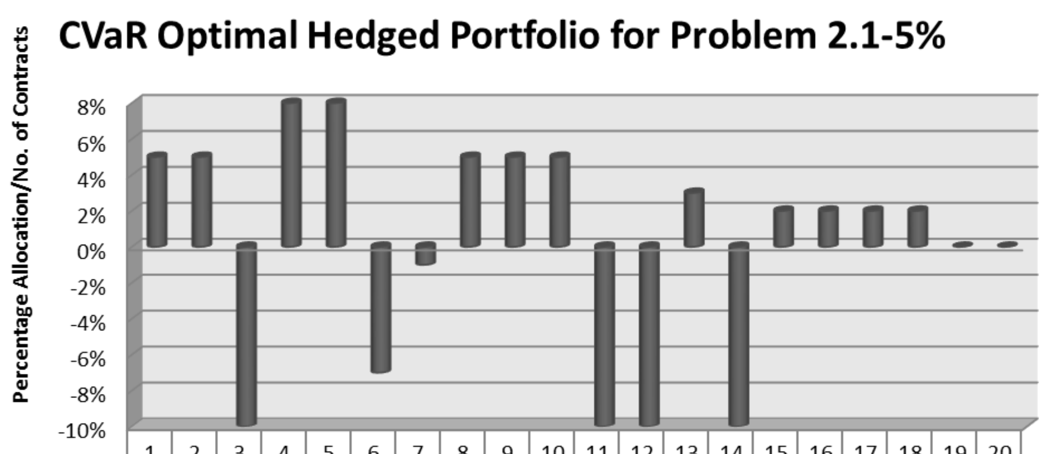
The hedging horizon t is the maturity of the barrier call option and the loss of the existing portfolio is,

$$\Pi^0(S,t) = P_0^{\text{int}} + \max(S_t - K, 0)\,. \qquad 7.15$$

The long barrier knock-out European call option shall be priced using a Monte Carlo method.

Results

We solved Problem 2.1 using an interior point algorithm in MATLAB for two different returns (5% and 10%). We notice that the different percentage allocations or the number of contracts for each asset, as well as a change from short to long positions for some assets (e.g. see asset 7). Solutions were also attempted using the active-set and simplex method. Both additional methods either yielded no solution or failed to converge adequately.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Prob 2.1-5%	5%	5%	-10%	8%	8%	-7%	-1%	5%	5%	5%	-10%	-10%	3%	-10%	2%	2%	2%	2%	0%	0%

Figure 7-13: CVaR optimal portfolio for Problem 2.1 with return =5%

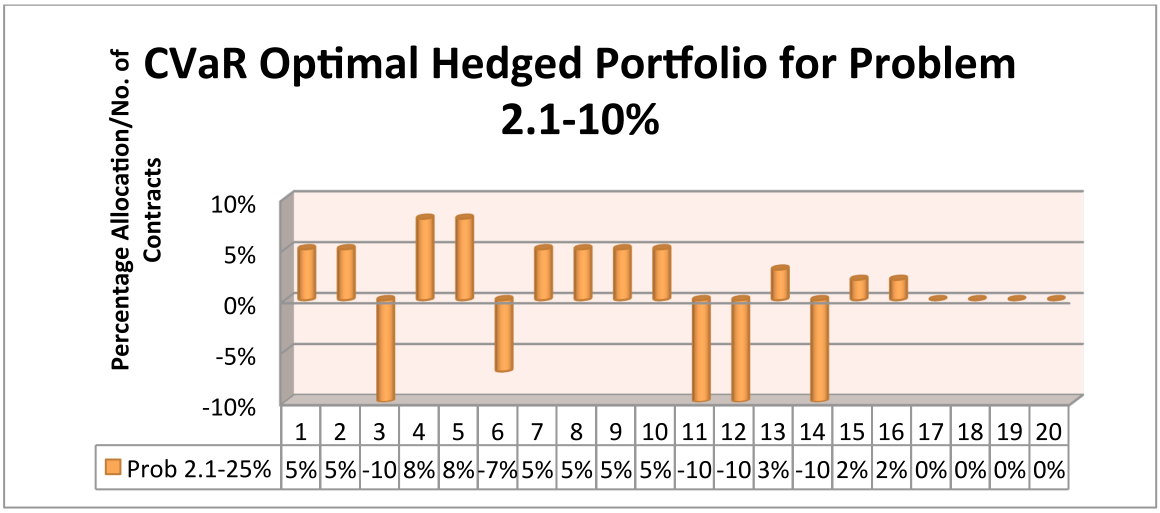


Figure 7-14: CVaR optimal portfolio for Problem 2.1 with return=10%

Figure *7-15* illustrates the optimal portfolio positions for the hedging problem involving the barrier option in Problem 2.2. We wished to hedge the barrier option in Problme 2.2 for short period of time , upto a year, with a portfolio (Dataset 2.4) with weights as given below. We notice that for a given different target portfolio, Problem 2.1 and Problem 2.2, and with the same hedge portfolio, a different optimal portfolio is obtained as compared to Figure *7-13* and Figure *7-14*. This confirms that there is some level of consistency with regards to the optimisation model setup.

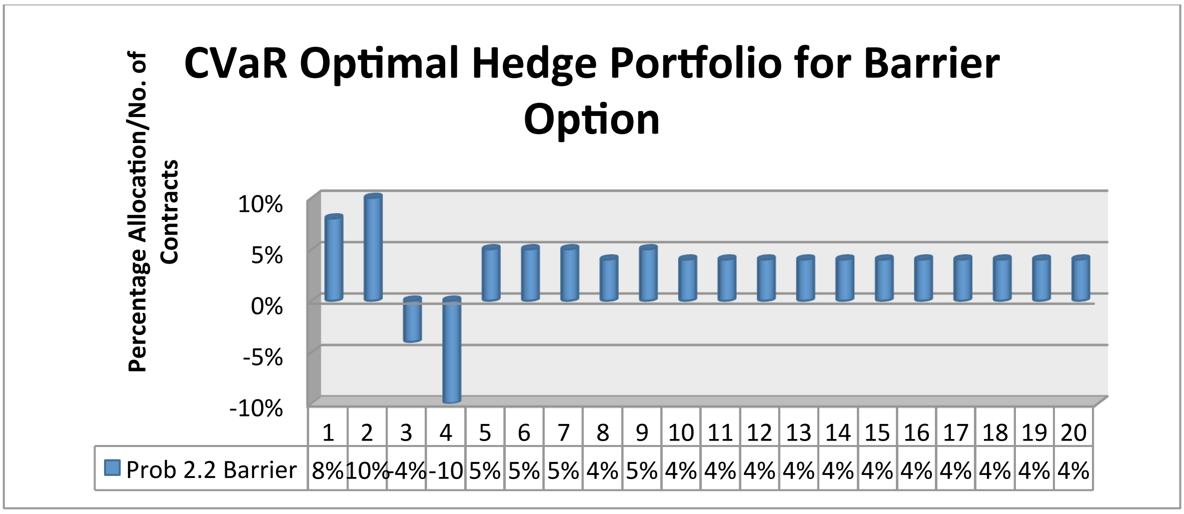


Figure 7-15: CVaR optimal hedge portfolio for barrier option

7.2.1 An investigation into time effects of hedging barrier options

In this investigation, we explore the effects of time on a barrier option to understand the risk and reward profile characteristics. We base the problem setup on the Problem 2.2 and Datasets 2.3. We solve for the optimal portfolio for different days to expiry. Algorithm 1 and 2 show the method used to analyse this characteristic.

Algorithm 1

---

m=Dataset 2.2

ZRisk(tdays,mu,m)

→*Begin*

→*Solve Barrier_Option(Dataset 2.3, tdays, nsim,n)*

→*Solve Initial Portfolio*

→*Derive f as per Eq. (4.59)*

→*Setup Linear Program*

→*Assign Aeq, beq, LB,UB,X0*

→*Assign optim_options*

→*Solve portfolio*

→*Get Risk for given tday and mu*

→*end*

Algorithm 2

Clear Variables

Initiate Variables

For return= 5% to 50 %

For tday= 5 days to end_days //e.g. end_days =100

→*Solve Z=ZRisk(Dataset, tday, return, m)*

end

end

Reshape (Z)

Plot(Z)

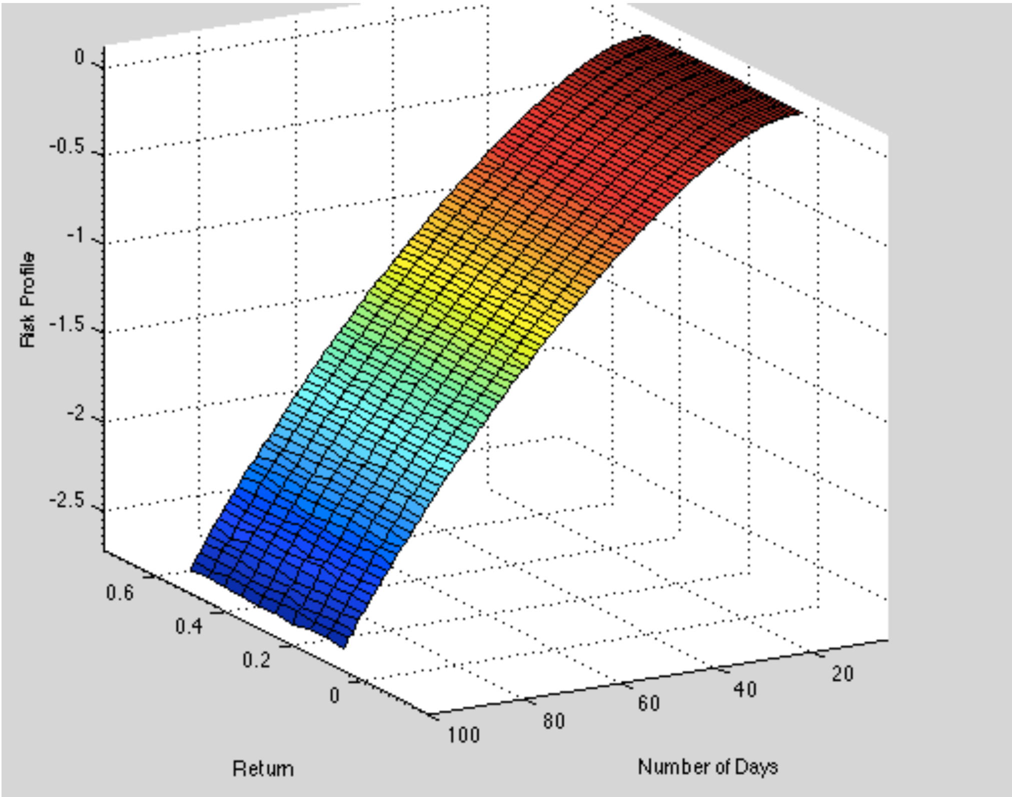


Figure 7-16: Risk profile plot showing effects of time and returns

Figure 7-16 illustrates the effects in a three-dimensional manner. We notice that as the barrier option approaches expiry it exhibits a concave characteristic to the risk profile. This typically means the hedging problem becomes less risky in a quadratic manner.

To understand the plot, in Figure 7-16 above, we need to understand some relationships that we have introduced in tandem with some commonly understood relationships. Since time is involved, the theta of a barrier option should be explored. Figure *7-17* below shows a plot of the option premium to the days to expiry of a barrier option. The figure illustrates the concave characteristics of time decay for an option.

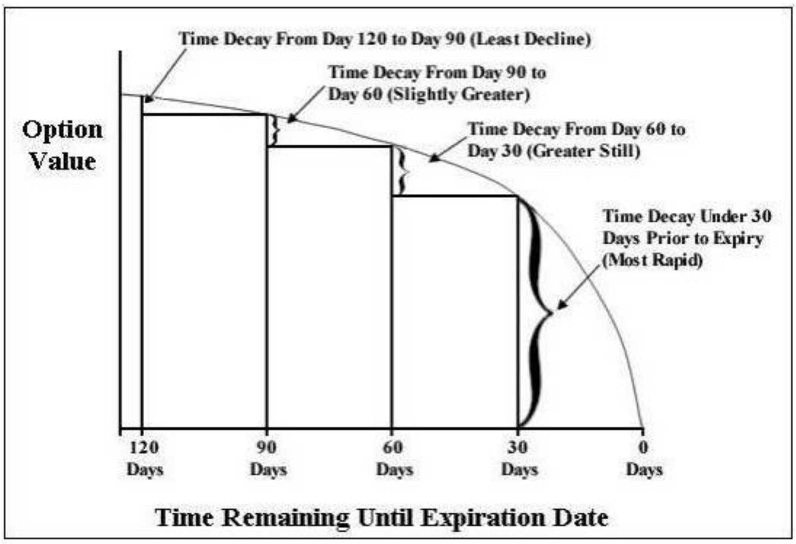


Figure 7-17: Relationship of option value to days till expiry [28]

In Eq. (7.16), we show the relationship between risk and change in a security value:

$$Risk = \alpha + \frac{1}{m(1-\beta)} \sum_{i=1}^{m} \left[-(\delta V_i)^T x - \alpha \right]. \qquad 7.16$$

We see that the relation between risk and change in security price has a directly proportional effect. The bigger the change more riskier the position becomes and vice versa. Comparing the theta for a barrier option and the risk to (δV_i) relationship can lead to an inverse relationship for risk and $\frac{1}{\delta t}$. So as time approaches expiry, this means that t of the option becomes larger (t=10, t=20, t=40) and as the time to expiry becomes smaller, your risk should become smaller. This reasoning offers an explanation of the concave plot illustrated in Figure 7-16.

Remarks

This case study is able to formulate and solve optimal hedging problems using the CVaR as a risk measure. We solved optimal hedging problems for short call options and long KO barrier options are considered as target portfolios. We also investigated the effects of time on a barrier option to understand the risk and reward profile characteristics. This case study achieves the goal set in Obj. (III).

7.3 Case Study 3: Portfolio optimisation with nonlinear transaction costs

This case study demonstrates a setup of a portfolio optimisation problem when risk is measured by CVaR with nonlinear costs. The problem dataset comprises of a set of options with various characteristics. This incudes the presence of different volatilities, European and American styles, and with various methods to solve the derivative pricing. Such methods may include analytic pricing, Monte Carlo and Delta-Gamma methods. The cost parameter shall be a weighted cost parameter on the optimal portfolio. The following weighted cost parameter shall be considered:

$$c_i = \omega \cdot \left|\overline{CVaR}\right| \qquad 7.17$$

where $\overline{CVaR}$ denotes the optimal CVaR with no cost consideration. See Section 4.2.4 for further details. We implicitly assume that the transaction costs of the instruments are the same.

Dataset 3.1

Let's suppose we have a portfolio that comprises of the following as given in Table 7-7 below. We to wish optimise this portfolio my minimizing the risk exposure.

Table 7-7: Table of information for Dataset 3.1

Asset	Type 1	Type 2	Asset Price	Asset Strike Price	Volatility	Interest Rate	t(days)
1	Call	European	60	50	0.1	0.05	60
2	Call	European	50	50	0.13	0.05	60
3	Call	European	80	50	0.14	0.05	60
4	Put	American	50	50	0.28	0.05	60
5	Call	European	50	50	0.1	0.05	60
6	Call	European	60	50	0.16	0.05	60
7	Put	American	50	50	0.1	0.05	60
8	Call	European	50	50	0.3	0.05	60
9	Call	European	50	50	0.1	0.05	60
10	Call	European	50	50	0.1	0.05	60
11	Put	European	75	50	0.2	0.05	60
12	Put	American	100	70	0.24	0.05	60
13	Put	European	75	50	0.1	0.05	60
14	Put	European	75	50	0.23	0.05	60
15	Put	European	75	50	0.1	0.05	60
16	Put	American	75	50	0.1	0.05	60
17	Put	European	100	75	0.21	0.05	60
18	Put	European	75	50	0.1	0.05	60
19	Put	European	75	50	0.22	0.05	60
20	Put	European	100	50	0.22	0.05	60

Problem 3.1

Problem 3.1 will use the portfolio given in Dataset 3.1 (which is the same Dataset 1.1) and aims to solve the following with cost considerations,

$$\min_{(x,\alpha)}\left(\alpha+\frac{1}{m(1-\beta)}\sum_{i=1}^{m}\left[-(\delta V_i)^T x-\alpha\right]+\sum_{j=1}^{n}c_j\left|x_j\right|\right)$$

Subject to:

$$\left(V^0\right)^T x=1, \qquad 7.18$$

$$\left(\overline{\delta V}\right)^T x=\bar{r},$$

$$l\le x\le u.$$

Note that if one models the cost as $\sum_{i=1}^{n}c^2{}_i x^2{}_i$ for example, the resulting optimal portfolio typically has few, if any, of its instruments with a small holding ratio $\left|x^*{}_i\right|$. For the quadratic penalty function, the constraint $x^*{}_i=0$ is only satisfied as the penalty parameter c_i tends to $+\infty$.

Results

Figure *7-18* below illustrates the portfolio weighting for each asset and the position of the asset (either short or long) for CVaR risk measure with cost considerations. We must compare the results in Figure *7-18* with Figure *7-1*, to see the effect of cost in the distribution of weights. Figure *7-18* below; illustrates more long positions as compared to Figure *7-1*. Additionally, the two soultions seem to solve for different problems and thus enhances the effects that cost have on portfolio optimisation. Costs become a very practical tool that the risk manager must consider to achieve for realistic portfolios. We must note that the interior point method solved this problem while the simplex method failed to converge.

The key advantage of CVaR over the spectral, distortion, CDRM and even the drawdown risk measure is that CVaR caters for cost in its optimisation model with much ease. The literature considered in this thesis has not found any cost consideration models for other risk measures and thus begs further research.

In order to analyse the impact of the cost consideration on risks, Alexander et al. [6] consider a relative difference of CVaR under different weighted cost parameters with repect to that under cost consideration. The formula can be represented as such:

$$\text{RelDifCVaR}(\omega) = \left| \frac{CVaR(\omega) - \left|\overline{CVaR}\right|}{\left|\overline{CVaR}\right|} \right|. \qquad 7.19$$

Based on Eq. (7.19), we shall present results for a single simulation problem and do not show averages. The results are tabulated in Table *7-8*.

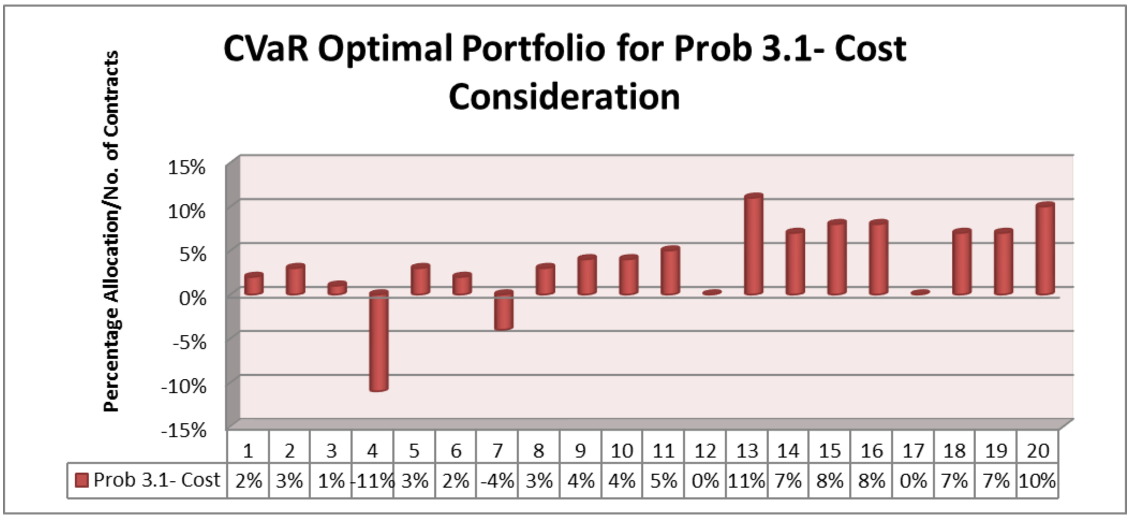


Figure 7-18: CVaR optimal Portfolio for 3.1 with cost consideration

Table *7-8* illustrates that by using CVaR and cost optimisation formulation, it is possible to obtain CVaR optimal portfolios with less number of securities with comparable risks. A similar result is also obtained in Alexander et al. [6]. We thus confirm the same overall result by using our data.

Table 7-8: Effect of weighted cost parameters on the optimal CVaR portfolio

β	ω	**CVaR**	**RelDifCVaR**	**# Securities**
0.95	0	1.655	0	20
	0.005	1.742	0.0321	15
	0.05	2.251	0.1170	8
0.99	0	1.699	0	20
	0.005	1.810	0.0331	15
	0.05	2.234	0.1202	8

One must also take into consideration the inevitable existence of model error and computational error.

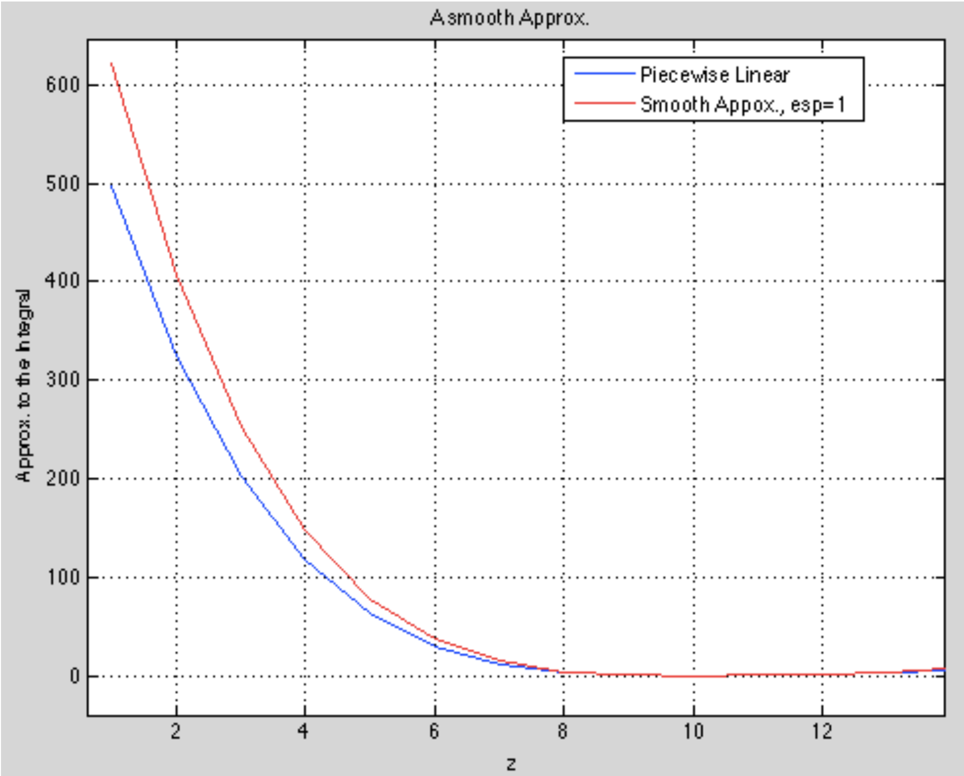


Figure 7-19: A smooth approximation plot

Figure *7-19* above shows a accurancy and smoothness of the CVaR optimisation problem. It uses the piecewise linear minimisation problem subject to linear constraints. We solve Eq. (7.18) by considering a continuously differentiable piecewise quadratic approximation given by:

$$\widetilde{F}_{\beta}(x,\alpha)=\left(\alpha+\frac{1}{m(1-\beta)}\sum_{i=1}^{m}\rho_{\varepsilon}\left[-(\delta V_i)^T x-\alpha\right]\right) \qquad 7.20$$

where $\rho_{\varepsilon}(z)$ is a continuously differentiable piecewise quadratic function which approximates the piecewise linear function max (z,0) given a resolution parameter $\varepsilon > 0$,

$$\rho_{\varepsilon}(z)=\begin{cases} z, & \text{if } z \geq \varepsilon. \\ \frac{z^2}{4\varepsilon}+0.5z+0.25\varepsilon, & \text{if } -\varepsilon \leq z \leq \varepsilon. \\ 0, & \text{otherwise.} \end{cases} \qquad 7.21$$

Table 7-9 below shows a simple test applied to the American style options on a MacBook Air, Mid 2012, 1.8 Ghz Core i5, 4Gb 1600Mhz DDR3. We compare the execution time with and without applying the smoothing technique.

Table 7-9: Matlab simulation times for with and without smoothing

Number of Monte Carlo simulations (American)	**Matlab- (CPU sec) without smoothing**	**Matlab- (CPU sec) with smoothing**	**Increase in efficiency (%)**
10000	0.32	0.11	65%
25000	0.72	0.56	22%
50000	0.96	0.78	19%

The comparison between the CPU times of the proposed smoothing formulation and the LP approach (interior point algorithm) for individual problems is illustrated in Table 7-9. The implementation of the smoothing method is based on an interior point method for nonlinear minimisation with bound constraints and is implemented in Matlab 13. The results above show the improvement in execution time, in seconds to execute with a dereasing efficiency for more number of Monte Carlo simulations. This means that the smoothing technique is much more efficient than the LP approach and one should explore using different efficiency techniques of Monte Carlo simulation in addtion to the smoothing techique. This thus confirms the work of Alexander et al. [6].

Remarks

In this case study, we effectively solved a CVaR based optimisation problem with cost considerations. We conducted a comparative analysis using a relative difference formula that compares the models risk with and without cost considerations. We concluded that with variations in the weighted cost scalar, ω we obtain a smaller portfolio number of securities with acceptable risk. We also included results where we used a smoothing technique to solve the model in a more efficient manner. This case study achieves the goal of objective Obj. (IV).

7.4 Case Study 4: Empirical risk analysis on the FTSE/JSE All-Share Index (ALSI)

In this case study, we perform an empirical risk analysis on the FTSE/JSE All-share index using the drawdown risk measure. The FTSE/JSE Index Series is designed to represent the performance of South African companies, providing investors with a comprehensive and complementary set of indices, which measure the performance of the major capital and industry segments of the South African market. The FTSE/JSE All-Share Index represents 99% of the full market capital value, i.e. before the application of any investability weightings, of all ordinary securities listed on the main board of the JSE, subject to minimum free-float and liquidity criteria. The objective of the index is for use in the creation of index tracking funds, derivatives, and as a performance benchmark.

Performance and Volatility - Total Return

Index (ZAR)	Return %						Return pa %*		Volatility %**		
	3M	6M	YTD	12M	3YR	5YR	3YR	5YR	1YR	3YR	5YR
FTSE/JSE All-Share	8.9	-0.2	5.6	0.2	49.6	87.9	14.4	13.4	18.4	15.1	11.1
FTSE/JSE Top 40	6.6	-2.5	2.6	-0.7	49.1	82.7	14.2	12.8	19.8	16.2	12.3

* Compound annual returns measured over 3 and 5 years respectively
** Volatility – 1YR based on 12 months daily data. 3YR based on weekly data (Wednesday to Wednesday). 5YR based on monthly data

Year-on-Year Performance - Total Return

Index % (ZAR)	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
FTSE/JSE All-Share	41.2	19.2	-23.2	32.1	19.0	2.6	26.7	21.4	10.9	5.1
FTSE/JSE Top 40	40.9	19.0	-23.6	31.7	17.2	2.2	26.1	22.8	9.2	7.5

Return/Risk Ratio and Drawdown - Total Return

Index (ZAR)	Return/Risk Ratio				Drawdown (%)			
	1YR	3YR	5YR	10YR	1YR	3YR	5YR	10YR
FTSE/JSE All-Share	0.0	1.0	1.2	0.8	-15.0	-15.0	-15.0	-45.4
FTSE/JSE Top 40	0.0	0.9	1.0	0.8	-15.1	-15.1	-15.1	-48.3

Return/Risk Ratio – based on compound annual returns and volatility in Performance and Volatility table
Drawdown - based on daily data

Figure 7-20: Performance information for FTSE/JSE ALSI from fact sheet

Figure *7-20* illustrates the data for th FTSE/ JSE ALSI. Figure *7-21* illustrates a ten-year view of the FTSE/JSE ALSI closing prices and the absolute drawdown series generated in Eq. (3.31). The datain Figure *7-20* is taken from the FTSE/JSE ALSI datasheet. Figure *7-21* illustrates the dip in prices during the 2008 financial crisis. In this case study, we shall show the following as part of the empirical risk analysis of the FTSE/JSE ALSI:

i. Fit a statistical distribution to the historic P&L distribution of the FTSE/JSE ALSI for 2015.

ii. Fit a statistical distribution to the historic P&L distribution of the FTSE/JSE ALSI for 2010-2016.

iii. Fit a statistical distribution to the drawdown series P&L distribution of the FTSE/JSE ALSI for 2010-2016.

iv. An investigation into the CVaR statistical analysis of P&L of the FTSE/JSE ALSI.

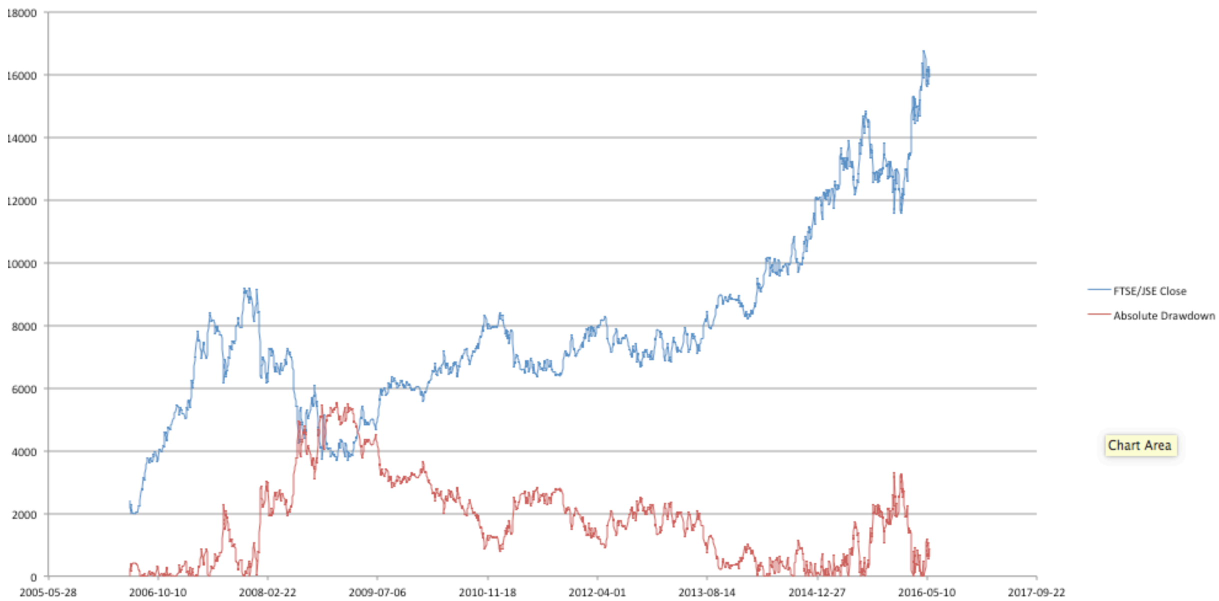


Figure 7-21: Plot of FTSE/JSE ALSI closing and the absolute drawdown

Fit a statistical distribution to the historic P&L distribution of the FTSE/JSE ALSI for 2015

We shall fit a statistical distribution to the historic P&L distribution of the FTSE/JSE ALSI for the period of 2015. These distributional fits where performed with Matlab's Statistical Toolbox-Distribution Fitting Application. Definitions of the distributions used, the reader is advised to refer to Prekopa [31].

Table 7-10: Distributional parameters for FTSE/JSE ALSI 2015 data

Distribution	**Parameters**	**Values**
Extreme Value	Mu	0.00850503

	Sigma	0.0206723
Generalized Pareto	Mean	-0.853297
	Variance	3.04078
	K	-1.7526
	Sigma	1.01209
	Theta	-0.5
Normal	Mu	0.000233894
	Sigma	0.0171265
Student-t	Mu	8.25331e-07
	Sigma	1.15656e-09
	nu	-1.10913e-06

Table *7-10* contains the statistical parameter values for each of the different statistical distributions considered. Along with Figure *7-22* and Figure *7-23*, illustrates that the distribution for historic P&L data of the FTSE/JSE ALSI for 2015 data displays that the distribution can be modelled with Student t or Normal distribution. The Student t paramenters are very low and should resemble a normal distribution. Matlab's-Statistical Toolbox- Distribution Fitting Application suggests that Student t fits more accurantly. The Generalised Pareto distribution does not capture the nature of distribution at all. The Generalize Extereme Values further do not catpure the nature as there are iverweighted tails and underweighted centers. Thus these two should not be considered in the best-fit modelling.

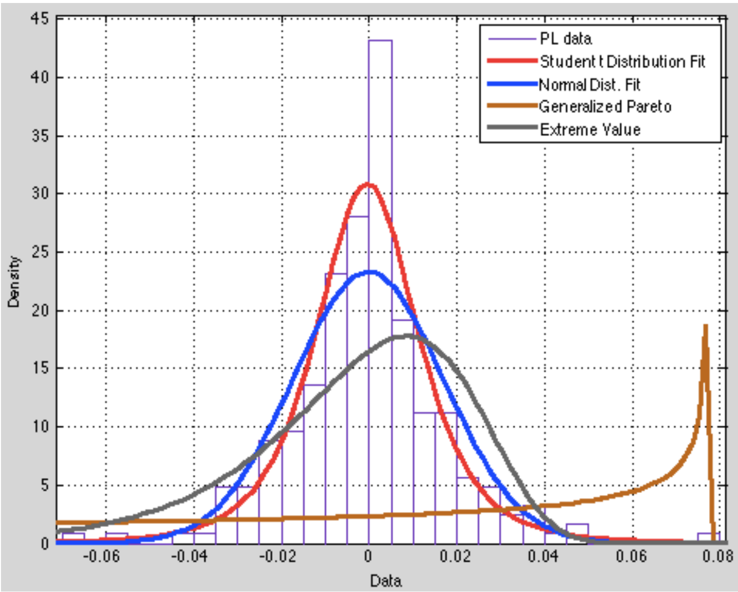


Figure 7-22: Distribution fitting for FTSE/JSE ALSI 2015 data

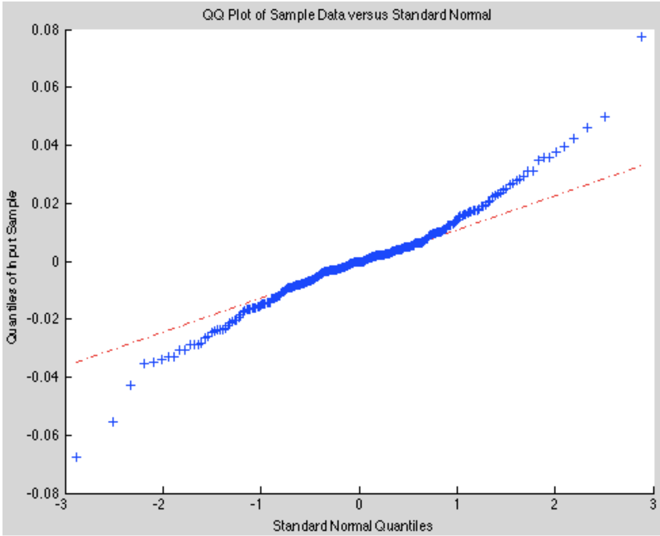


Figure 7-23: Q-Q plot for FTSE/JSE ALSI 2015 data

Fit a statistical distribution to the historic P&L distribution for the FTSE/JSE ALSI for 2010-2016

We shall fit a statistical distribution to the historic P&L distribution of the FTSE/JSE ALSI for the period of 2010 to 2016. These distributional fits where performed with Matlab's Statistical Toolbox- Distribution Fitting Application.

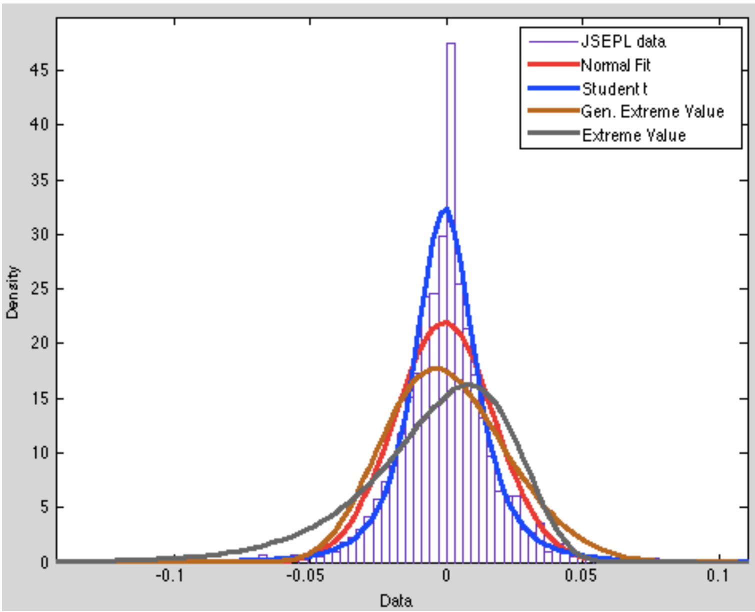


Figure 7-24: Distributional fit for FTSE/JSE ALSI 2010-2015 data

Table *7-11* contains the statistical parameter values for each of the different statistical distributions considered. Along with Figure *7-24* and Figure *7-25*, we notice that the distribution for FTSE/JSE ALSI for the 2010 to 2015 data displays that the distribution can be modelled better with the Student t distribution. These distributional fits were performed with Matlab's Statistical Toolbox- Distribution Fitting Application.

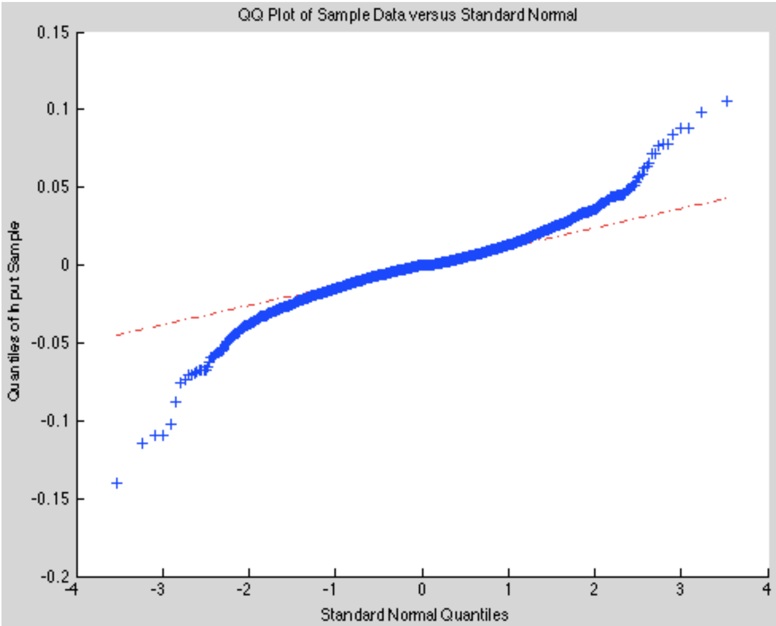


Figure 7-25: Q-Q plot for FTSE/JSE ALSI 2010-2015 data

Table 7-11: Distributional parameters for FTSE/JSE ALSI 2010-2015 data

Distribution	**Parameters**	**Values**
Extreme Value	Mu	0.00828215
	Sigma	0.0226998
Generalized Extreme Value	Mean	0.000754341
	Variance	0.000506742
	K	-0.183423
	Sigma	0.0211458
	mu	-0.00815249
Normal	Mu	-0.000763004
	Sigma	0.0181756
Student-t	Mu	-0.000626607
	Sigma	0.0113254
	nu	2.91806

Table *7-12* contains the statistical parameter values for each of the different statistical distributions considered. Along with Figure *7-26*, we notice that the distribution for FTSE/JSE ALSI for the 2010 to 2015 drawdown data displays that the distribution can be modelled better with the Generalised Pareto distribution. These distributional fits were performed with Matlab's-Statistical Toolbox- Distribution Fitting Application.

Fit a statistical distribution to the drawdown series P&L distribution for FTSE/JSE ALSI for 2010-2016

We shall fit a statistical distribution to the drawdown series P&L distribution of the FTSE/JSE ALSI for the period of 2010 to 2016. These distributional fits were performed with Matlab's Statistical Toolbox- Distribution Fitting Application.

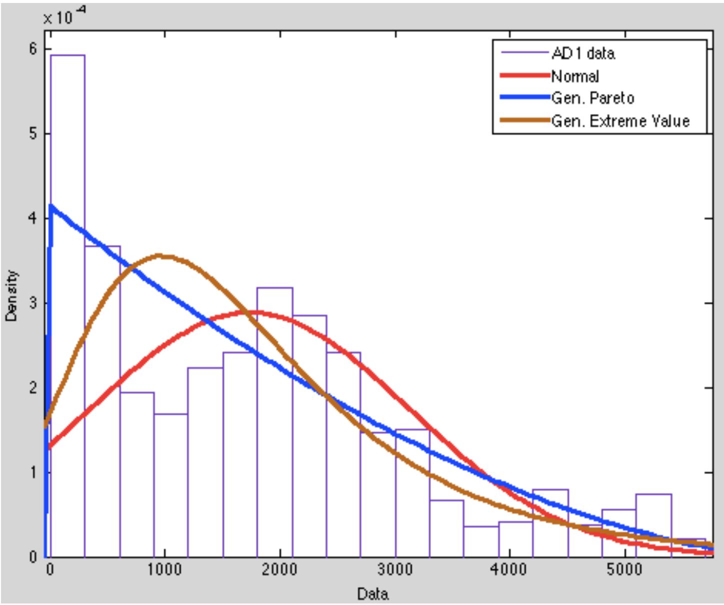


Figure 7-26: Distributional fit for absolute drawdown series of FTSE/JSE ALSI 2010-2015 data

Table 7-12: Distributional parameters for FTSE/JSE ALSI 2010- 2015 absolute drawdown data

Distribution	**Parameters**	**Values**
Generalized Pareto	Mean	713.962
	Variance	5.34501e+06
	K	-0.374463
	Sigma	2414.36
	mu	-2
Generalized Extreme Value	Mean	1736.12
	Variance	2.24668e+06
	K	0.080074
	Sigma	1039.02

	mu	1047.48
Normal	Mu	1733.99
	Sigma	1383.85

Table 7-13: Calculated and data from fact sheet comparison of FTSE/JSE ALSI risk measure data

	1YR	3YR	5YR
Data from Fact Sheet, Drawdown Measure	-15.0%	-15.0%	-15.0%
Calculated Drawdown Measure	-13.4%	-15.1%	-14.3%
Calculated CVaR	-15.5%	-15.5%	-15.8%

Table *7-13* contains comparative infomation pertaining to the three risk measures based from the fact sheet and calculated either with the drawdown measure or CVaR measure. We notice that the calculated drawdown measure is aligned with that obtained from the fact sheet while we see that the CVaR measure forms higher values of risk as compared to the drawdown measure. This confirms the commonly understood notion that CVaR does cater for a larger loss of information as compared to VaR and now the drawdown risk measure.

An investigation into the CVaR statistical analysis of FTSE/JSE ALSI

In Section 6.4.1, we presented a statistical analysis of CVaR. The work comprises of the findings from Stoyanov et al. [38]. We presented the CVaR and tail thickness sensitivity formulae from [38]. From the findings of Case Study 4, we showed that the Profit/Loss of FTSE/JSE ALSI follows a Student t distribution. In this investigation we wish to determine if the findings of [38] can be replicated using the FTSE/JSE ALSI statistical data.

Set up of the data used in the investigation:

i. We have taken the historical data for the FTSE/ JSE ALSI from Jan 2000 to Dec 2015. (approx. 2400 data points).
ii. We then use a 200-day rolling sequence to the data, practically 200 trading days are considered. Due to rolling of the existing data set (2400 points), we then develop the relevant dataset of 42200 data points.
iii. We then work out the volatility (σ), Drift (μ) and the degrees of freedom (υ) for each of the datasets generated in (ii).

iv. Based on the calculation in (iii), we then approximate the fat-tail analysis equations for the Student t distribution based on Eq. (6.8), Eq (6.9) and Eq (6.10). The approximate equations are as follows:

$$\frac{\partial CVaR_{\varepsilon}(\upsilon,\sigma,\mu)}{\partial \upsilon} \approx \frac{\Delta CVaR_{\varepsilon}(\upsilon,\sigma,\mu)_{datasets}}{\Delta \upsilon_{datasets}} \approx -\sigma \frac{\partial CVaR_{\varepsilon}(\upsilon,1,0)}{\partial \upsilon}, \qquad 7.22$$

$$\frac{\partial CVaR_{\varepsilon}(\upsilon,\sigma,\mu)}{\partial \sigma} \approx \frac{\Delta CVaR_{\varepsilon}(\upsilon,\sigma,\mu)_{datasets}}{\Delta \sigma_{datasets}} \approx CVaR_{\varepsilon}(\upsilon,1,0), \qquad 7.23$$

$$\frac{\partial CVaR_{\varepsilon}(\upsilon,\sigma,\mu)}{\partial \mu} \approx \frac{\Delta CVaR_{\varepsilon}(\upsilon,\sigma,\mu)_{datasets}}{\Delta \mu_{datasets}} \approx -1. \qquad 7.24$$

v. We then plot the distributions and find how well they approximate the analytic calculations obtained in Section 6.4.1 by the work of Stoyanov et al. [38]

In order to achieve the results of Eq. (7.22), we need to normalize $CVaR_{\varepsilon}(\upsilon,\sigma,\mu)$ to $CVaR_{\varepsilon}(\upsilon,1,0)$. We can achieve this by performing the following:

$$CVaR_{\varepsilon}(\upsilon,\sigma,\mu) = \frac{CVaR_{\varepsilon}(\upsilon,\sigma,\mu) - \mu}{\sigma} \qquad 7.25$$

We then approximate the following as:

$$\begin{gathered} \frac{\partial CVaR_{\varepsilon}(\upsilon,\sigma,\mu)}{\partial \upsilon} \approx \frac{\Delta CVaR_{\varepsilon}(\upsilon,\sigma,\mu)_{datasets}}{\Delta \upsilon_{datasets}}, \\ -\sigma \frac{\partial CVaR_{\varepsilon}(\upsilon,1,0)}{\partial \upsilon} \approx -\sigma \frac{\Delta CVaR_{\varepsilon}(\upsilon,1,0)_{datasets}}{\Delta \upsilon_{datasets}}. \end{gathered} \qquad 7.26$$

After performing this data manipulation, we super-impose and compare each result. We see this result in Figure *7-27* below. We see that there is close fit of the data and we can safely conclude that the data analysis conforms to the analytic derivations of Stoyanov et al. [38] for Eq(7.27).

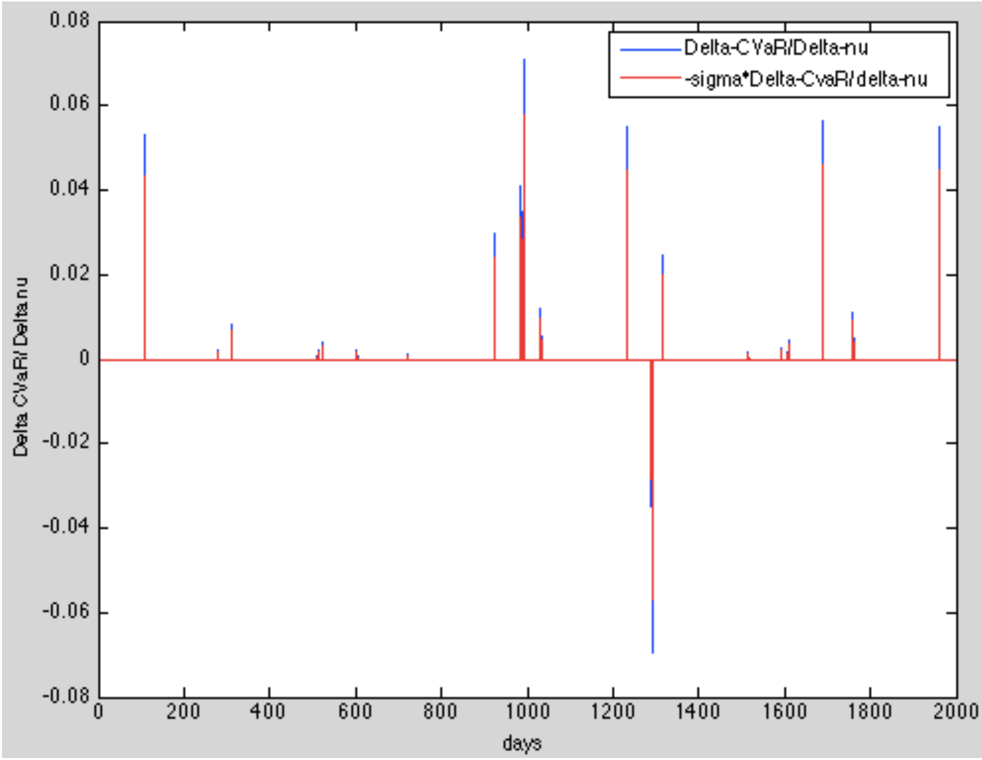


Figure 7-27: Superimposed plot for Eq. (7.27)

For the second relational equation Eq. (7.28),we need to normalize $CVaR_{\varepsilon}(\upsilon,\sigma,\mu)$ to $CVaR_{\varepsilon}(\upsilon,1,0)$. This is shown in Eq. (7.25). Figure *7-28* below shows the plot of $CVaR_{\varepsilon}(\upsilon,1,0)$ and the plot of $\frac{\partial CVaR_{\varepsilon}(\upsilon,\sigma,\mu)}{\partial\sigma} \approx \frac{\Delta CVaR_{\varepsilon}(\upsilon,\sigma,\mu)_{datasets}}{\Delta\sigma_{datasets}}$ superimposed. See that the error is of small magnitude with 94% correlation.

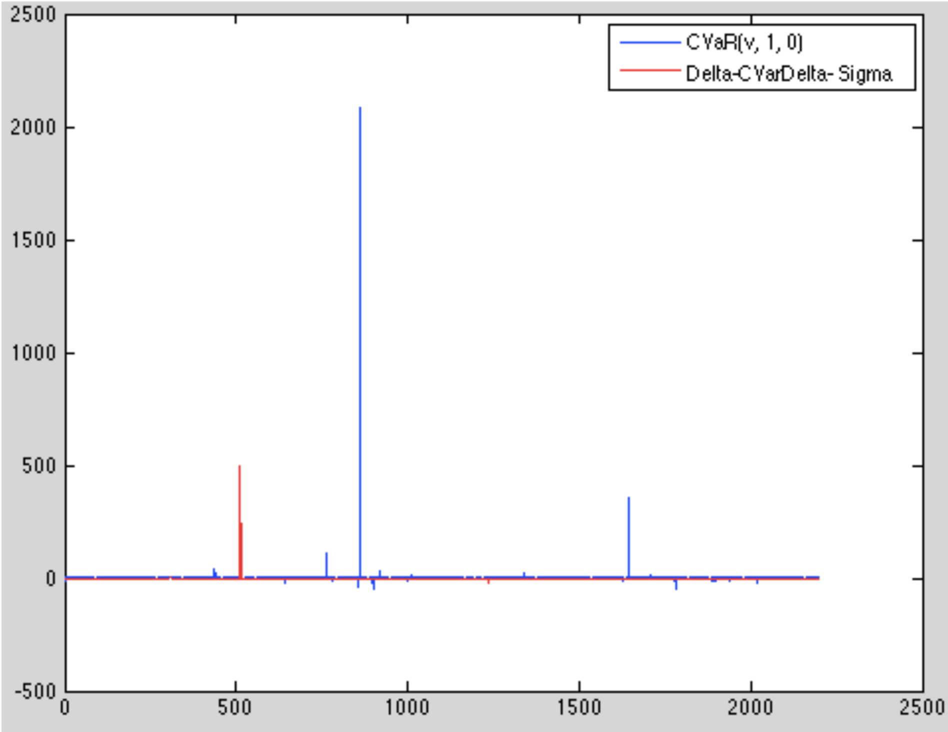


Figure 7-28: Superimposed plot for Eq. (7.28)

For Eq. (7.29) above, we simply plot the data analysis and observe that the overall trend of the result is -1 with some minor deviations, this is illustrated in Figure 7-29 . We can conclude that the data analysis confirms the analytic result obtained from Stoyanov et al. [38].

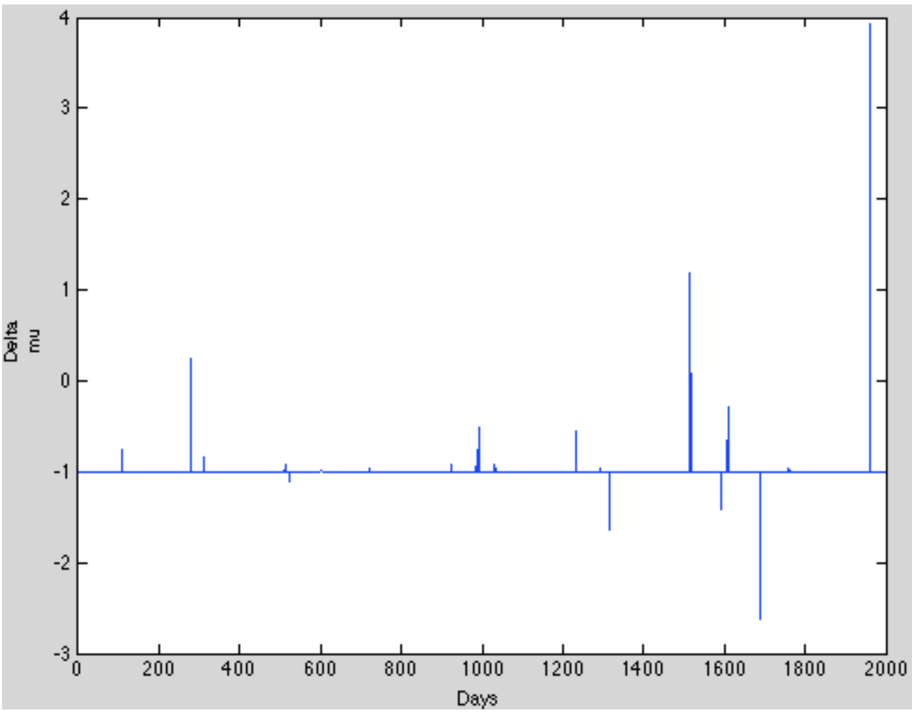


Figure 7-29: Plot for Eq. (7.29)

This study confirms our investigation into the statistical properties of CVaR. We confirm the analytic derived equations and relations derived from Stoyanov et al. [38] with the data analysis. Until enough information is provided to make conclusions, however the errors can be attributed to the model and statistical variations.

Remarks

In this case study we did an emperical analysis of the FTSE/JSE ALSI historical data. We also did statistical fitting of data for the absolute drawdown series of the FTSE/JSE ALSI P&L data. We showed what is already known by so many risk managers, that assuming normal distibrutions to data is erroneous and full distributional modelling must be performed to get more accurate results. In the analysis of the FTSE/JSE ALSI, we showed Student t distribution is more prominent than the normal distribution. We also confirmed the derivatives of CVaR with respect to the distributional parameter by using a data analysis to confirm the relation equations of Stoyanov et al. [38] in Section 6.4.1.

7.5 Case Study 5: Portfolio optimisation with drawdown risk measure

This case study demonstrates an optimisation setup for Conditional Drawdown (CDD). We shall solve for a portfolio optimisation problem with a static set of weights using the drawdown measure. In a hedge fund company, there are multiple traded assets on the futures markets. Such hedge funds use complicated mathematical strategies for executing such strategies, which give them a certain edge.

Most of the hedge fund traders trade so-called long-term trend-following systems, but there are now multiple examples of short-term mean reverting trading systems as well. These systems may be viewed as functions of the individual futures market price realized prior to the present time. These strategies normally have a substantial smoothing-out effect on the futures prices and have close to stationary properties. Every hedge fund, has to allocate a certain portion of overall risk (or overall capital that it manages) to each and every "market". Due to a substantial level of stationarity of the strategies, each hedge fund calculates the weights according to a certain internal proprietary weight allocation procedure. Normally, this set remains fixed and does not change unless a certain market gets added or removed from the set, which normally happens when a new system is introduced, when a certain market disappears or a new market is being added. We shall use the historical data for constructing the optimal static portfolio using drawdown as the risk measure. Below is the list of markets used in this case study. For more informtion regarding typical hedge funds, the reader is advised to read Chekhlov et al. [11].

Table 7-14: List of markets for Case Study 5

Asset No.	Ticker	Asset description
1	AAO	The Australian All Ordinaries Index
2	AD	Australian Dollar Currency Futures
3	BD	U.S. Long (30- Year) Treasury Bond Futures
4	CD	Canadian Dollar Currency Futures
5	CP	Copper Futures
6	DX	U.S. Dollar Index Currency Futures
7	ED	90-Day Euro Dollar Futures
8	FV	U.S. 5 Year Treasury Note Futures
9	JY	Japanese Yen Currency Futures
10	LFT	FTSE-100 Index Futures

The data consists of a time series that covers five years of data ranging from 2010 to 2015. Time is measured in trading days only, with a standard of five workdays per week. Problems 5.1 and 5.2 resemble the static asset allocation problems performed by Chekhlov et al. [11]. We consider a smaller set of 10 assets instead of the 32 assets Chekhlov et al. [11] considered. Problem 5.1 consists of solving a static portfolio using the maximum drawdown risk measure. Problem 5.2 consists of solving a the same static portfolio of Problem 5.1 but by using the average drawdown risk measure instead. We shall compare results from Problem 5.1 and Problem 5.2 by using CVaR as risk measure in the static portfolio optimisation problem.

Problem 5.1

In problem 5.1 we shall construct the static portfolio given in Table *7-14*. Based on Section 5.2.1, the reduction to linear program given as:

$$
\begin{aligned}
&\max_{u,y,z} \sum_{j=1}^{K} p_j w_{jN}(x) \\
&\text{Subject to}, \\
&\sum_{i=1}^{L} \chi_i \left(y + \frac{1}{(1-\alpha)N} \sum_{k=1}^{N} z_{ik} \right) \le \gamma, \\
&z_{ik} \ge u_{jk} - y_i, \\
&u_{jk} \ge u_{jk-j} - r_{jk}, \\
&u_{j0} = 0, \\
&u_{jk} \ge 0, \\
&z_{ijk} \ge 0, \\
&i = 1,...,L, \\
&j = 1,...,K, \\
&k = 1,...,N.
\end{aligned}
\qquad 7.27
$$

Where u_{jk}, y_i and z_{ijk} are auxiliary variables. We solve the following problem with the maximum drawdown risk measure as a constraint with the objective function being the rate of return:

$$
\begin{aligned}
&\max_{u,y,z} \sum_{j=1}^{K} p_j w_{jN}(x) \\
&\text{Subject to,} \\
&u_{jk} \geq u_{jk-j} - r_{jk}, \\
&\gamma \geq u_{jk} \geq 0, \\
&u_{j0} = 0, \\
&z_{ijk} \geq 0, \\
&j = 1,...,K, \\
&k = 1,...,N.
\end{aligned}
\qquad 7.28
$$

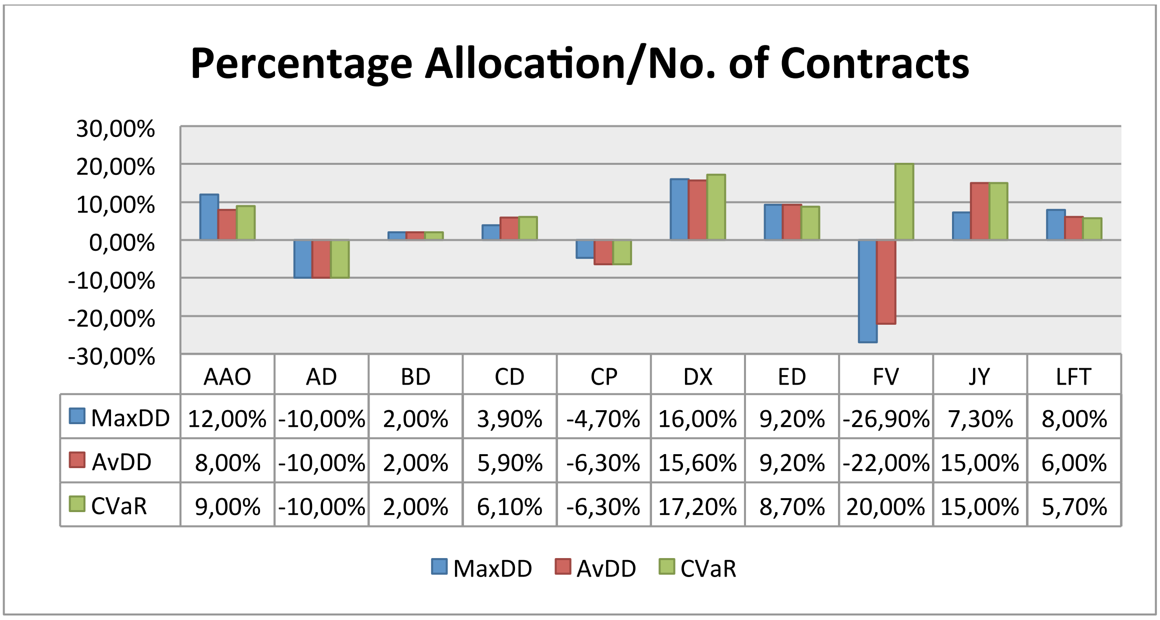
Problem 5.2

In problem 5.2 we shall construct a LP model to solve the static portfolio given in Table *7-14*. We solve the following problem with the average drawdown risk measure as a constraint with the objective function being the rate of return:

$$
\begin{aligned}
&\max_{u,y,z} \sum_{j=1}^{K} p_j w_{jN}(x) \\
&\text{Subject to,} \\
&u_{jk} \geq u_{jk-j} - r_{jk}, \\
&\frac{1}{N} \sum_{k=1}^{N} \sum_{j=1}^{K} p_j u_{jk} \leq \gamma, \\
&u_{jk} \geq 0, \\
&u_{j0} = 0, \\
&z_{ijk} \geq 0, \\
&j = 1,...,K, \\
&k = 1,...,N.
\end{aligned}
\qquad 7.29
$$

Results

We shall now consider the combined results of the Problem 5.1 and Problem 5.2. Figure *7-30* illustrates the percentage allocation or the number of contracts to be considered for creating an optimal portfolio based on using the maximum and average drawdown measures and CVaR. Chekhlov et al. [11] do not state the methods they used to solve the knapsack problem in Eq. (7.28) and Eq. (7.29), but we considered using a genetic algorithm that solved the optimisation problem programmed and solved with Matlab. Other methods that can be used can be found in the work of Ermoliev and Wets [14] and Pflug [30].



	AAO	AD	BD	CD	CP	DX	ED	FV	JY	LFT
MaxDD	12,00%	-10,00%	2,00%	3,90%	-4,70%	16,00%	9,20%	-26,90%	7,30%	8,00%
AvDD	8,00%	-10,00%	2,00%	5,90%	-6,30%	15,60%	9,20%	-22,00%	15,00%	6,00%
CVaR	9,00%	-10,00%	2,00%	6,10%	-6,30%	17,20%	8,70%	20,00%	15,00%	5,70%

Figure 7-30: Percentage allocation/ No. of contracts for Case Study 5

Figure *7-31* illustrates the efficient frontier constructed for the three measures. We notice that both the maximum and average drawdown measures show the concave nature of the common understood relationship of return and risk based on Markowitz [24].

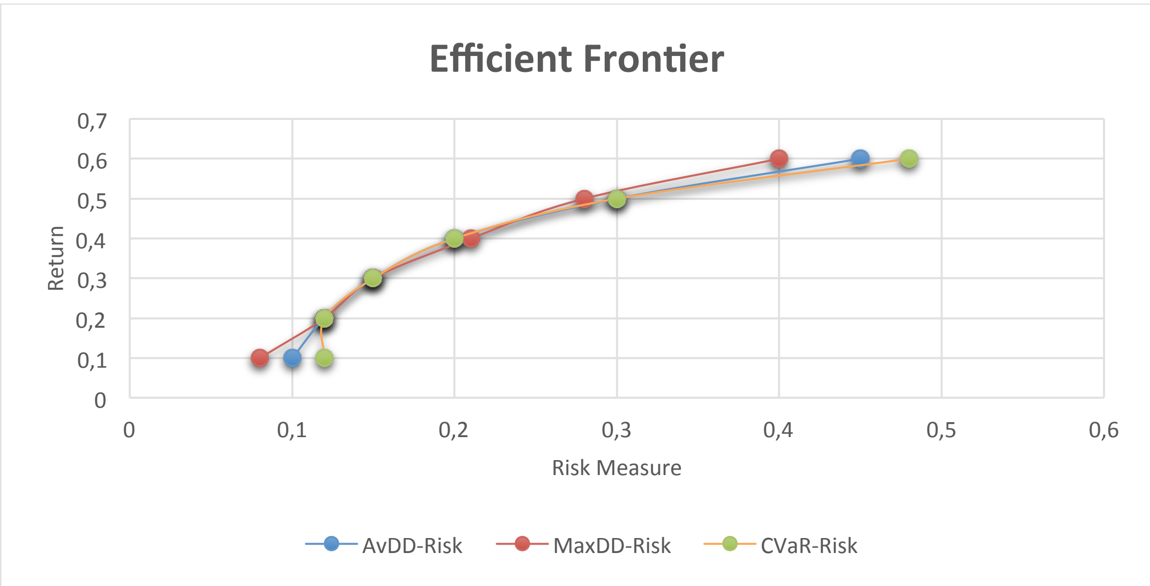


Figure 7-31: Efficient frontier plot for Case Study 5

Chekhlov et al. [11] consider using "box constraints" on the portfolio weights. We also adopt this and include the constraint of $2\% \leq x_i \leq 80\%$. Based on Chekhlov et al. [11], in a futures setup constraint are analogous to the "fully-invested" condition from the original work of Markowitz [24], this makes the efficient frontier strictly concave. Figure *7-31* illustrates this characteristic. Without the constraint on the portfolio weights, the efficient frontier would be a straight line passing through the origin (0,0). Chekhlov et al. [11] state that is due to the virtually infinite leverage of these types of strategies.

If we consider our case, we see that for a lower bound of 2%, 10x0.02=0.2 as minimal leverage while the upper case of 80%, 10x0.8=8 as maximal leverage. The optimal allocation of weights chooses both the optimal leverage and allocation proportions between instruments. Chekhlov et al. [11] considered another very important issue that has to do with the stability of the optimal portfolio. They note that the constraints need to lead to sufficiently stable portfolios by providing enough mixing of the individual equity curves [11].

Remarks

In this case study we solveed a static porfolio using the maximum and average drawdown risk measures and compared the results by using the CVaR risk measure. We have used a genetic algorithm to solve the knapsack problems presented in Problem 5.1 and Problem 5.2. This case study acheives the goal set out in objective Obj (V).

Chapter 8. Conclusions

In this chapter, we present some concluding remarks on this thesis. We present a summary of the work and the important findings and results. We shall relate each of the objectives set out in Section 1.3 to what we have achieved. In section 1.3, we presented some of the main objectives we wish to acheive in this thesis. A summary of the objectives is:

(Obj. I). We shall present the theorems, properties and other propositions for the Value-at-Risk (VaR), Conditional Value-at-Risk (CVaR), spectral risk measures, distortion risk measures and coherent distortion risk measures from various literature resources. We shall use the Artzner et al. [8] axiomatic framework to form the basis of each risk measure's definition.

(Obj. II). We shall present for each of the risk measures above, an optimisation model(s) that can be solved to give the risk manager an optimal portfolio. The model solution will be based on risk or return or both as an objective function and constraint. These optimal problems shall be formulated for derivative based assets.

a. We wish to investigate the effects of solving a portfolio optimisation problem with each of the three risk measures, namely, CVaR, spectral risk measures and CDRM risk measure. These three risk measure have shown to have some popularity in both literature and in practice thus a comparative analysis among them may prove insightful.

(Obj. III). We shall also present the use of the CVaR risk measure to formulate an optimal hedging problem with derivative based contracts. This problem is typically constructed with a given target portfolio and a given hedging portfolio. The aim is to hedge the target portfolio with the given hedging portfolio. This problem also lends itself well to the practical setting of a risk manager who wishes to hedge risk of his current portfolio(s).

(Obj. IV). We shall present a CVaR risk measure optimisation model that considers transactional or managerial costs in the objective function. This problem is essentially covered by Alexander et al. [6], where they formulate the CVaR optimal portfolio model. We shall use their work to form the optimal problem and solve it for our given portfolio(s).

(Obj. V). Chekhlov et al. [11] introduced a new one-parameter of family risk measures called Conditional Drawdown (CDD). We shall present the theorems and, properties of this risk measure. We then formulate an optimal portfolio problem that shall be solved using the Conditional Drawdown (CDD) as either objective function or constraint. These problems generally lead to knapsack type optimisation problems and we shall use genetic algorithms to solve the problems.

We shall now present the key findings and outcomes for each of the objectives that was set out in Chapter 1 of this document.

In Chapter 2 and 3, we presented a number of risk measures. Some of the key risk measures that we used in this thesis are CVaR, spectral risk measures, distortion risk measures, CDRM, and conditional drawdown risk measure. An axiomatic approach to formalising each risk measures definition and key properties were shown.

We presented the distortion risk measure. It possessed some key desirable properties for a portfolio risk measure, i.e. law-invariance, sub-additivity, and consistency with the second order stochastic dominance. In addition, distortion risk measures have their roots in the distortion utility theory of choice under uncertainty, meaning that this class of risk measures can better reflect the risk preferences of investors.

Spectral risk measures were also presented with their theoretical and practical properties. We showed how to formulate an optimal portfolio model using the spectral sequence and estimator of a spectral measure. Various spectrums exist and we considered three variants when solving for the spectral risk measure based portfolio optimisation in Case Study 1.

We presented the CDRM which is an extended version of the well-known linear optimisation framework for CVaR to a general class of risk measure. We presented the finite generation theorem for CDRM in Bertsimas and Brown [8] and showed that any CDRM can be defined as a convex combination of ordered portfolio losses and equivalently a convex combination of CVaRs. Based on this we presented a CDRM based portfolio optimisation model. We solved a CDRM-based portfolio optimisation via linear programming, which could handle problems with a large number of variables and/or constraints.

We presented the drawdown risk measure, which has proved useful for practical portfolio management [11]. This measure similar to CVaR, includes the MaxDD and AvDD measures as its limiting cases and possesses all properties of a deviation measure [11]. Moreover, it may be considered as a generalization of deviation measure to a dynamic case. We considered optimisation models that addressed an asset-allocation problem with CDD, MaxDD and AvDD measures. In case study 5, we formulated a real life optimisation problem. We solved the knapsack optimisation problem using a genetic algorithm.

In Chapter 4, we looked at CVaR as a loss distribution and then showed how we can use CVaR in portfolio optimisation. We presented the model for using CVaR as an objective function or as a constraint. We presented the inclusion of a weighted cost consideration in the CVaR optimisation problem. The model represented cost as proportional to the magnitude of instrument holding. This cost model is capable of controlling transaction cost as well as management cost. We illustrated that minimizing CVaR together with this cost model leads to more desirable portfolios with significantly smaller transaction costs, fewer non-zero instrument holdings, and comparable CVaR measures. We presented a computationally efficient method for solving a simulation based CVaR optimisation problem by exploiting the fact that the objective function in the CVaR optimisation problem approaches a continuously differentiable function as the number of Monte Carlo samples increases.

We presented the investigations and scenarios for creating and solving optimal portfolios in a case study format. We presented five case studies each achieving the same global outcome of using risk measures in the application of portfolio optimisation. We shall now give an account of further learnings from each case study.

In Case Study 1, we solved the portfolio optimisation problem by finding optimal weights of derivative securities using CVaR, spectral and CDRM as risk measures. In section 4, we developed a comprehensive model that solves an optimal portfolio using CVaR as a risk measure. Similarly, in Section 5 we developed the portfolio optimisation models for spectral, CDRM and conditional drawdown (CDD) risk measures. We constructed the efficient frontiers for each risk measure model and performed a comparative analysis. We have noted that from a trader's perspective, the CVaR offered the best fit of the traditional risk-reward characteristic while some deficiencies where observed for the spectral and CDRM risk measures.

The spectral risk measure efficient frontier illustrated some "flatness" for lower risk values. This led us to do some further investigation. We included two more spectrums and then solved the model, and this also yielded the same "flatness" for lower risk values. Based on this, we then setup another simple portfolio of stocks only (no derivatives) and optimised it using the spectral risk measures. The results of the efficient frontier on the simple stock portfolio yielded a more concave plot of risk and reward. This led to the conclusion that the "flatness" is due to some characteristics of derivatives that yielded these effects for lower risk values. This poses the question: why derivative based portfolio optimisation using spectral risk measures exhibit such characteristics on the efficient frontiers for lower risk values? We shall leave this for further investigative research.

From a comparative perspective, we used the efficient frontier as a mechanism for comparing CVaR, spectral and CDRM risk measures in portfolio optimisation. Based on the comparative efficient frontier plot, Figure *7-10*, we see that for lower levels of risk CVaR and CDRM show similarities while at higher risk levels CVaR and spectral risk measures show some similarities. Based on the risk-return comparative method, spectral and CRDM cause some difficulties specially from a trader's perspective. As a topic for further research, a quatitative method to rank the different risk measures needs to be developed.

In Case Study 2 , we were able to formulate and solve optimal hedging problems using the CVaR as a risk measure. We solved optimal hedging problems for short call options and long KO barrier options as target portfolios. We also investigated the effects of time on a barrier option to understand the risk and reward profile characteristics. This case study achieves the goal set in objective Obj. (III).

In Case Study 3, we effectively solved a CVaR based optimisation problem with cost considerations. We conducted a comparative analysis using a relative difference formula that compares the model's risk with and without cost considerations. We concluded that with variations in the weighted cost scalar, ω , we obtain smaller number of securities in the portfolio with acceptable risk. We also included results where we used a smoothing technique to solve the model in a more efficient manner. This case study achieves the goal set in objective Obj. (IV).

In Case Study 4, we did an emperical analysis of the FTSE/JSE ALSI historical data. We also performed a statistical fitting of data for the absolute drawdown series of the FTSE/JSE ALSI P&L data. We showed what is already known to many risk managers, that assuming normal distibrutions to data is erroneous and full distributional modelling must be performed to obtain more accurate results. In the analysis of the FTSE/JSE ALSI, we have showed that the Student t distribution is more prominent than the normal distribution. We also confirmed the derivatives of CVaR with respect to the distributional parameter by using data analysis to confirm the relation equations of Stoyanov et al. [38] in Section 6.4.1.

In Case Study 5, we solved a static porfolio using the maximum and average drawdown risk measures and compared the results using the CVaR risk measure. We used a genetic algorithm to solve the knapsack problems presented in Problem 5.1 and Problem 5.2. This case study acheives the goal set out in objective Obj (V).

The objective of this thesis was to investigate risk measures and the application thereof to portfolio optimisation. We have performed an in depth look into the theoretical aspects of the risk measures within an axiomatic framework. We developed the portfolio optimisation models for the key risk measures and set up case studies that looked at the application details.

References

1. Acerbi, C., “Spectral measures of risk: A coherent representation of subjective risk aversion”, Journal f Banking and Finance 26, 1505-1518, 2002
2. Acerbi, C., Simonetti P., “Portfolio Optimisation with Spectral measure of Risk”, March 2002
3. Adam, A., Houkari, M., Laurent J.P., “Spectral risk measures and portfolio selection”, Journal of Banking & Finance 32, 1870-1882, 2008
4. Adam, A., Houkari, M., Laurent, JP., “Spectral risk Measures and portfolio Section”, Journal of Banking and Finance 32, p.p. 1870-1882, 2008
5. Alexander, S., Coleman, T.F., Li, Y. “Derivative Portfolio Hedging Based on CVaR”, New Risk Measures in Investments and Regulation by Szego, D., 2003
6. Alexander, S., Coleman, T.F., Li, Y. “Minimizing CVaR and VaR for a Portfolio of Derivatives”, Journal of Banking and Finance 30, pp 583-605, 2006
7. Artzner et al., “Coherent Multiperiod Risk-Adjusted Values and Bellman Principle”, Preprint, ETH Zurich, 2004
8. Artzner, P., Eber, F., Eber, J.M., Heath, D., “Coherent measures of risk”, Mathematical Finance 9, 203–228, 1999
9. Artzner, P., Eber, F., Eber, J.M., Heath, D., “Thinking Coherently”, Risk 10, 1997
10. Birge, J. R., Louveaux, F., “Introduction to Stochastic Programming”, Springer, 1997
11. Chekhlov A., Uryasev S., Zabarankin M., “Drawdown measure in portfolio optimisation”, Risk management and financial engineering lab, centre for Applied optimisation, Department of Industrial and systems engineering, June 2003
12. Dowd, K., Cotter, J., “Spectral Risk Measures with applications to Futures Clearinghouse Variation Margin Requirements”, Science direct, October 2006
13. Einmal, J. H. J., et al., “VaR Stress tests for highly non-linear portfolios”, Journal of Risk Finance, JRF 6,5, 2005
14. Ermoliev, Y., Wets, R., “Numerical Techniques for Stochastic Optimisation”, Springer Series in Computational Mathematics, Vol. 10, Springer, 1988
15. Feng, M. B., Tan K. S., “Coherent Distortion Risk Measures in Portfolio Selection”, Systems Engineering Procedia 2, 25-34, 2012
16. Follmer, H., Penner, I., “Convex risk measures and the dynamics of their penalty functions”, Stat. Decisions 24(1),61-96, 2006
17. Follmer, H., Schied, A., “Convex and Coherent risk measure”, Finance Stoch. 6, 429–447, 2008
18. Follmer, H., Schied, A., “Convex measures of risk and trading constraints”, Finance Stoch. 6, 429–447, 2002
19. Goldberg L., Mahmoud O., “On a convex measure of drawdown risk”, department of statistics and economics and centre for risk management research, University of California, Berkeley, September 2015
20. Hans-Jakob, L., Doege, J., “Convex Risk Measures for Portfolio Optimisation and concepts of Flexibility”, ETH, Zurich-Swiss Federal Institute of Technology, 26 Jan 2005
21. Kall, P., Wallace, S.W., “Stochastic Programming”, Wiley, 1995
22. Kan, Y.S., Kibzun, A.I.,” Stochastic Programming Problems with Probability and Quantile functions”, Wiley, 1996
23. Longstaff, F.A., Schwartz, E. S., “Valuing American Options by Simulation: A Simple Least-Squares Approach”, The Review of Financial Studies, Vol 14, No 1, pp. 113-147., 2001
24. Markowitz, H., “Portfolio Selection”, Journal of Finance 7 (1), 77-91, 1952

25. Matmoura, Y., Penev, S., “Multistage optimisation of option portfolio using higher order coherent risk measure”, European Journal of Operations Research 227, 190-198, 2013
26. Mausser, H., Rosen, D., “Beyond VaR: From Measuring Risk to Managing Risk”, ALGO research Quarterly 1, 5-20
27. McKay, R., Keefer, T.E., “VaR is a dangerous technique”, Corporate Finance. Searching for Systems Integration Supplement, 1996
28. P. GLASSERMAN, “Monte Carlo Methods in Financial Engineering”, Springer, Applications in Mathematics, Second Edition, 2003
29. Pflug, G. Ch., “Some remarks on value at risk and conditional value at risk”, Uryasev, S. (Ed.), Probabilistic constrained: Methodology and Applications, Kulwer Academic Publishers, Dordrecht, 2000
30. Pflug, G.Ch., “Optimisation of Stochastic Models: The Interface Between Simulation and Optimisation”, Kluwer, 1996
31. Prekopa, A. “Stochastic Programming”, Kluwer, 1995
32. Rachev, S. T. et al., “Distortion Risk Measures in Portfolio Optimisation”, Handbook of portfolio construction, 2010
33. Rachev, Z., Loteva, B. R., Stoyanov, S., “Capturing fat tails”, Risk Magazine, May 2010
34. Rachev, Z., Loteva, B. R., Stoyanov, S., Fabozzi, F. J., “Fat-tailed models for risk estimation”, Working paper series in Economics, May 2011
35. Rockafellar, R.T., Uryasev, S., “Conditional value-at-risk for general loss distributions”, Journal of Banking and Finance 26 (7), 1443–1471, 2006
36. Rockafellar, R.T., Uryasev, S., “Optimisation of conditional value-at-risk”, Journal of Risk 2 (3), 21– 41, 2000
37. Stoyanov, S., Rachev, S. T., Fabozzi, F. J., “CVaR sensitivity with respect to tail thickness”, March 2011
38. Stoyanov, S., Rachev, S. T., Fabozzi, F. J., “Sensitivity of portfolio VaR and CVaR to portfolio return characteristics”, May 2011
39. Uryasev, S., “Conditional value-at-risk: Optimisation algorithms and applications”, Financial Engineering News, 14, 2000
40. Wang S., Young V., Panjar H., “Axiomatic characterization of insurance prices”, Insurance: Mathematics and Economics 21, pp 337-347, 1997

Appendix

A1. Further theorems and proof for convex measures of risk

Theorem A1.1:

Suppose $\mathrm{X} = L^{\infty}(\Omega, A, P)$, P is a set of probability measures $Q \ll P$, and $\rho : X \to \mathbb{R}$ is a convex measure of risk. Then the following properties are equivalent.

There is a "penalty function" $\alpha : \mathrm{P} \to (-\infty, \infty)$ such that

$$\rho(Y) \geq \sup_{Q \ll \mathrm{P}} \left(E_Q[-X] - \alpha(Q) \right) \text{ for all } X \in \mathrm{X} \qquad 0.1$$

The acceptance set A_{ρ} associated with ρ is weak, i.e. $\sigma\left(L^{\infty}(P), L^{1}(P)\right)$-closed ρ possesses the Fatou property: if the sequence $(X_n)_{n\in\mathbb{N}} \subset \mathrm{X}$ is uniformly bounded and X_n converges to some $X \in \mathrm{X}$ in probability, then $\rho(X) \leq \liminf_n \rho(X_n)$

if the sequence $(X_n)_{n\in\mathbb{N}} \subset \mathrm{X}$ decreases to $X \in \mathrm{X}$, then $\rho(X_n) \to \rho(X)$

Proof of Theorem A1.1

1 implies 2 holds, because ρ given by (16) is $\sigma\left(L^{\infty}(P), L^{1}(P)\right)$-lower semi-continuous. For the converse implication, we can repeat the proof of Theorem 5 and apply the Hahn-Banach separation theorem in the locally convex space $\left(L^{\infty}(P), \sigma\left(L^{\infty}(P), L^{1}(P)\right)\right)$ in order to get a negative continuous linear functional ℓ satisfying (14). By assumption, ℓ can be represented as $\ell(Z) = E[\varphi X]$ with some $\varphi \in L^{1}(P)$ yielding a probability measure $dQ/dP = \varphi / E[\varphi]$. We conclude the proof as in Theorem 5. The remaining implications follow as in [4].

Proposition A1.1:

Suppose $\rho : L^{\infty}(\Omega, A, P) \to \mathbb{R}$ is convex measure of risk possessing a representation of the form (16) and take P as in Theorem 6. Then the representation (16) holds as well in terms of the penalty function

$$\alpha_0(Q) = \sup_{X \in L^{\infty}} \left(E_Q[-X] - \rho(X) \right) = \sup_{X \in A_{\rho}} E_Q[-X]$$

Moreover, is minimal in the sense that $\alpha_0(Q) \leq \alpha(Q)$ for all $Q \in P$ if the representation (16) holds for $\alpha(\bullet)$. In addition,

$$\alpha_0(Q) = \sup_{X \in A_\rho} E_Q[-X] = \sup_{X \in A} E_Q[-X]$$

if ρ is defined as in (8) via a given acceptance set A.

We may have a penalty function α_0 that is not the minimal one and this may occur in the following case presented in Proposition A1.2.

Proposition A1.2

Suppose that for every i in some index set I we are given a convex measure of risk ρ_i on $X =: L^\infty(\Omega, A, P)$ with associated penalty function $\alpha_i(\bullet)$. We assume that

$$\inf_{Q \in \wp} \inf_{i \in I} \alpha_i(Q) > -\infty \qquad 0.2$$

Then

$$\rho(X) := \sup_{i \in I} \rho_i(X), \qquad X \in \mathrm{X} \qquad 0.3$$

is a convex measure of risk that can be represented as (16) with the penalty function

$$\alpha(Q) := \inf_{i \in I} \alpha_i(Q), \qquad Q \ll P \qquad 0.4$$

Proof:

Clearly:

$$\rho(X) = \sup_{i \in I} \sup_{Q \ll P} \Big(E_Q[-X] - \alpha_i(X) \Big) = \sup_{Q \ll P} \Big(E_Q[-X] - \inf_{i \in I} \alpha_i(Q) \Big) \qquad 0.5$$

Hence the assertion follows.

A.2 Convex efficient frontier for conditional drawdown risk

Now we shall present the logic behind the above statement. Denoting $g(x)=\sum_{j=1}^{K} p_j w_{jN}(x)$, we show that for any $\gamma_{1,2}\in[0,1]$ and $\tau\in[0,1]$

$$g\left(x_\chi^*\left(\tau\gamma_1+(1-\tau)\gamma_2\right)\right)\geq \tau g\left(x_\chi^*\left(\gamma_1+(1-\tau)g\left(x_\chi^*\left(\gamma_2\right)\right)\right)\right) \qquad 0.6$$

According to Eq. (5.15), we have

$$\begin{aligned} & g\left(x_\chi^*(\gamma)\right)=\max_{x\in X,y} g(x) \\ & s.t.\ H(x,y)\leq\gamma \end{aligned} \qquad 0.7$$

and using notation $G_\lambda(x,y)=g(x)-\lambda H(x,y)$, we obtain

$$g\left(x_\chi^*(\gamma)\right)=\min_{\lambda\geq 0}\max_{x\in X,y}\left(G_\lambda(x,y)+\lambda\gamma\right)=\min_{\lambda\geq 0}\left(G_\lambda\left(x(\lambda),y(\lambda)\right)+\lambda\gamma\right) \qquad 0.8$$

Since expression $G_\lambda\left(x(\lambda),y(\lambda)\right)+\lambda\gamma$ is linear with respect to γ, $\min_{\lambda\geq 0}\left(G_\lambda\left(x(\lambda),y(\lambda)\right)+\lambda\gamma\right)$ is a concave function of γ. Indeed,

$$\begin{aligned} & \min_{\lambda\geq 0}\left(G_\lambda\left(x(\lambda),y(\lambda)\right)+\lambda\left(\tau\gamma_1+(1-\tau)\gamma_2\right)\right) \\ & =\min_{\lambda\geq 0}\left(\tau\left(G_\lambda\left(x(\lambda),y(\lambda)\right)+\lambda\gamma_1\right)+(1-\tau)\left(G_\lambda\left(x(\lambda),y(\lambda)\right)+\lambda\gamma_2\right)\right) \\ & =\tau\min_{\lambda\geq 0}\left(G_\lambda\left(x(\lambda),y(\lambda)\right)+\lambda\gamma_1\right)+(1-\tau)\min_{\lambda\geq 0}\left(G_\lambda\left(x(\lambda),y(\lambda)\right)+\lambda\gamma_2\right) \end{aligned} \qquad 0.9$$

A3. Numerical example using Longstaff and Schwartz method

For this example, we shall use 7 paths and 3 time periods.

Step 1: The simulated stock prices are shown below.

0	1	2	3
100	102.89	114.05	86.04
100	196.49	181.21	200.98
100	167.15	156.66	208.60
100	73.63	94.93	100.58
100	176.23	224.46	265.28
100	111.08	141.68	157.44
100	94.88	104.26	94.26

Step 2: Table below shows the values of the convertible at $t = 3$. The table shows 5 conversions and 2 redemptions at face value. These cash values will now be rolled back to $t = 2$

3	
86.04	Redemption
200.98	Conversion
208.60	Conversion
100.58	Conversion
265.28	Conversion
157.44	Conversion
94.26	Redemption

Step 3: The convertible will not be exercised in path 4 at $t = 2$. The value is =97.61 which the discount value of 94.91.

Step 4: the remaining time paths will need to continuation value to be calculated. We shall discount the cash flows paid out in $t = 3$ to $t = 2$

Step 5: use the present values to regress the continuation values to find the coefficients a, b, c from the basis function.

Path	S (2)	P*	Pc
1	114.05	114.05	110.66
2	181.21	181.21	207.08
3	156.66	156.66	180.31
4	94.93	94.93	97.61

5	224.46	224.46	225.87
6	141.68	141.68	160.47
7	104.26	104.26	87.39

Step 6: Then compare exercise and continuation values to generate intermediate cash flows at $t = 2$. The following cash flow matrix is obtained. Go to Step 3. And decrease time period to $t = 1$.

Path	**1**	**2**	**3**
1		114.05	
2			200.98
3			208.60
4		97.61	
5			265.28
6			157.44
7		104.26	

Step 3: at time, two points are suboptimal path 4 and 7. Calculate the continuation value for these path nodes:

Step 4: Calculate the present value for subsequent nodes 1, 2, 3, 5, 6. The calculation is shown Table below.

Step 5: Use regression and generate continuation values.

Step 6: Compare P* and Pc and determine the value of the convertible bond producing Table below:

0	**S (1)**	**P***	**Pc**	**P**	
100	102.89	102.89	117.71	117.71	
100	196.49	196.49	194.61	194.61	Convert
100	167.15	167.15	222.78	222.78	
100	73.63	73.63	94.73	94.73	
100	176.23	176.23	221.03	221.03	
100	111.08	111.08	138.33	138.33	
100	94.88	94.88	101.18	101.18	

Evaluation of Convertible bond has the final cash flow matrix shown below.

Path	**1**	**2**	**3**

1		114.05	
2	196.49		200.98
3			208.60
4	94.73	97.61	
5			265.28
6			157.44
7	101.18	104.26	

Discounting the cash values to the present time yields a value of =152.16.

A3. FTSE/JSE ALSI additional information

Top 10 Constituents

Constituent	ICB Sector	Net MCap (ZARm)	Wgt %
SABMiller	Beverages	852,254	11.68
Naspers	Media	838,063	11.49
Compagnie Financiere Richemont AG	Personal Goods	474,260	6.50
BHP Billiton	Mining	411,854	5.65
British American Tobacco PLC	Tobacco	264,865	3.63
MTN Group	Mobile Telecommunications	261,732	3.59
Sasol	Chemicals	256,704	3.52
Steinhoff International Holdings N.V.	Household Goods & Home Construction	226,658	3.11
Anglo American	Mining	209,399	2.87
Old Mutual	Life Insurance/Assurance	183,720	2.52
Totals		**3,979,509**	**54.55**

ICB Supersector Breakdown

ICB Code	ICB Supersector	No. of Cons	Net MCap (ZARm)	Wgt %
1300	Chemicals	4	278,494	3.82
1700	Basic Resources	19	1,117,054	15.31
2300	Construction & Materials	7	27,614	0.38
2700	Industrial Goods & Services	17	363,245	4.98
3300	Automobiles & Parts	1	3,357	0.05
3500	Food & Beverage	13	1,010,737	13.86
3700	Personal & Household Goods	3	965,783	13.24
4500	Health Care	7	277,817	3.81
5300	Retail	16	402,503	5.52
5500	Media	2	840,909	11.53
5700	Travel & Leisure	5	36,811	0.50
6500	Telecommunications	4	339,305	4.65
8300	Banks	7	444,199	6.09
8500	Insurance	7	392,893	5.39
8600	Real Estate	30	457,135	6.27
8700	Financial Services	19	312,368	4.28
9500	Technology	2	24,407	0.33
Totals		**163**	**7,294,632**	**100.00**

Index Characteristics

Attributes	FTSE/JSE All-Share
Number of constituents	163
Net MCap (ZARm)	7,294,632
Dividend Yield %	2.86
Constituent Sizes (Net MCap ZARm)	
Average	44,752
Largest	852,254
Smallest	497
Median	10,692
Weight of Largest Constituent (%)	11.68
Top 10 Holdings (% Index MCap)	54.55

A.4 Matlab programs

Case Study 1

```
%This code solves for the simple case where CVAR is minimized with simple
%constraints. See the paper by Alexendra, Coleman, Li. This code
%corresponds to the Case Study 1
%Done By Resham Sivnarain, 11230712

%Matlab function descriptions
% X = linprog(f,A,b,Aeq,beq,LB,UB) defines a set of lower and upper
%     bounds on the design variables, X, so that the solution is in
%     the range LB <= X <= UB. Use empty matrices for LB and UB
%     if no bounds exist. Set LB(i) = -Inf if X(i) is unbounded below;
%     set UB(i) = Inf if X(i) is unbounded above.

% X = linprog(f,A,b) attempts to solve the linear programming problem:
%
%               min f'*x    subject to:   A*x <= b
%                x
%
%     X = linprog(f,A,b,Aeq,beq) solves the problem above while additionally
%     satisfying the equality constraints Aeq*x = beq.



clear ret;
clear risk;
for j=1:6

alp=0.05;
beta=0.95;
ret(j)=0.01*j;


[V0,dV]=analyticBS(m,Type1,Type2);


a=size(m);
f=-dV-alp;
A=zeros(1,a(1,1));
b=0;
Aeq=[V0;dV]
beq=[1; ret(j)]
LB=-0.3*ones(1,a(1,1));
UB=0.8*ones(1,a(1,1));
X=linprog(f,A,b,Aeq,beq,LB,UB)
risk(j)=f*X;
end

plot(risk,ret,'-o')
```

```
%%


%%%%%%%%%%%%%%%%
%X-CVAR PROBELM
%Y-SPECTRAL RISK PROBLEM
%Z-CDRM PORBLEM


alp=0.000;
beta=0.95;
retx=0.56;
rety=0.26;
retz=0.06;

[V0,dV]=analyticBS(m,Type1,Type2);


a=size(m);
f=-dV-alp;
A=zeros(1,a(1,1));
b=0;
Aeq=[V0;dV]
beqx=[1; retx]
beqy=[1; rety]
beqz=[1; retz]
LB=-1*ones(1,a(1,1));
UB=1*ones(1,a(1,1));
% X=linprog(f,A,b,Aeq,beq)
[X,FVALX, EXITFLAGX, OUTPUTX,LAMDAX]=linprog(f,A,b,Aeq,beqx,LB,UB);
[X,FVALY, EXITFLAGY, OUTPUTY,LAMDAY]=linprog(f,A,b,Aeq,beqy,LB,UB);
[X,FVALZ, EXITFLAGZ, OUTPUTZ,LAMDAZ]=linprog(f,A,b,Aeq,beqz,LB,UB);
x=100*X
Y=100*Y
Z=100*Z
% bar(X,'bl')
% hold on
% bar(Y,'r')
% hold on
% bar(Z,'c')


%%

N=100;
k=100;

for j=1:10
    q(j)=genspecseq(N,k,j);

end
plot(q)
```

```
%%

k=3;

for j=1:10
    f(j)=(j-1)/10;
    q(j)=specexp(k,f(j));


end
 plot(f,q,'r-o')

    hold on
k=15;

for j=1:10
    f(j)=(j-1)/10;
    q(j)=specexp(k,f(j));

end
plot(f,q,'b-x')
    hold on
% plot(f,q,'b')
```

```
function [Price,optstop,optcash]=American(S0,drift,vol,K,R,T,path,step)
%This function prices the American Options by using A Simple Least-Squares
%Approach developed by Longstaff and Schwartz
%Done by: Resham Sivnarain

% drift=0.0005;
% vol=0.05;

%Monte Carlo Simulation of the Stock Price
for j=1:path
    S(j,1)=S0;
for i=2:step
    S(j,i)=S(j,i-1)*exp(drift+randn(1)*vol);
end
end
%plots of the different MC paths
% for j=1:100
% plot(S(j,:));
% hold on
% end

% S=[1,1.09,1.08,1.34;...
%     1,1.16,1.26,1.54;...
%     1,1.22,1.07,1.03;...
```

```
%     1,0.93,0.97,0.92;...
%     1,1.11,1.56,1.52;...
%     1,.76,.77,0.9;...
%     1,0.92,0.84,1.01;...
%     1,0.88,1.22,1.34];

a=size(S);
r=a(1,1);
c=a(1,2);
% Calculate the discount rate-d
d=exp(-R*T);
%Optimal stopping Matrix
stopM=zeros(a(1:1),a(1,2));
%Cash flow Matrix
cfM=zeros(a(1:1),a(1,2));

%update the Cash flow Matrix to EU at final time
for i=1:r
    cfM(i,end)=max(0,K-S(i,end));
    if (cfM(i,end)>0)
        stopM(i,end)=1;
    end
end
cfM;
i=c-1;
for k=1:(c-2)
    X=S(:,i);
    Xn=S(:,i+1);
    t=zeros(r,1);
    cnt=1;
    for j=1:r
        E(j)=max(0,K-S(j,i));

        if (E(j)~=0)
            t(j)=1;                %tracking vector
            Xreg(cnt)=X(j);        %get Regression X
            Yreg(cnt)=cfM(j,i+1)*d; %get Regression Y
            cfM(j,i)=E(j);
            cnt=cnt+1;
        elseif (E(j)==0)
            X(j)=0;
        end
    end


p=mylsp(Yreg,Xreg);                  %work the regression constants

for j=1:(cnt-1)
    cont(j)=p(1)+p(2)*Xreg(j)+p(3)*(Xreg(j))^2; %pack the Vector
end
cnt=1;
for j=1:r
```

```
    if (t(j)==1)
        temp(j)=cont(cnt);
        cnt=cnt+1;
    elseif (t(j)~=1)
        temp(j)=0;
    end
end

cont=temp;
for j=1:r
    if (E(j)>cont(j))
        stopM(j,i)=1;
    end
end
i=i-1;
end

stopM;
cfM;
optstop=zeros(r,c);
kp=zeros(r,1);
for i=2:c
  for j=1:r
      if (stopM(j,i)==1) && (kp(j)==0)
          optstop(j,i)=stopM(j,i);
            kp(j)=1;
      end
  end

optcash=optstop.*cfM;
sum=0;
for i=1:r
    for j=1:c
        sum=sum+optcash(i,j)*exp(-R*(j-1));
    end
end

Price=sum/r;

end
```

Case Study 2

```
%This code solves for the hedging problem.
%constraints. See the paper by Alexendra, Coleman, Li. This code
%corresponds to the Case Study 2,
%Done By Resham Sivnarain, 11230712

%%
%Matlab function descriptions
```

```
% X = linprog(f,A,b,Aeq,beq,LB,UB) defines a set of lower and upper
%      bounds on the design variables, X, so that the solution is in
%      the range LB <= X <= UB. Use empty matrices for LB and UB
%      if no bounds exist. Set LB(i) = -Inf if X(i) is unbounded below;
%      set UB(i) = Inf if X(i) is unbounded above.

% X = linprog(f,A,b) attempts to solve the linear programming problem:
%
%               min f'*x    subject to:   A*x <= b
%                x
%
%      X = linprog(f,A,b,Aeq,beq) solves the problem above while additionally
%      satisfying the equality constraints Aeq*x = beq.


% [Call, Put] = blsprice(Price, Strike, Rate, Time, Volatility, Yield) computes European put
and call option prices using a Black-Scholes model.
clear ret;
clear risk;
ret=0.25;
alp=0.0;
beta=0.99;
mm=1;
w=0.005;

Init_port=blsprice(100,100,0.04, 10, 0.2,0.1)
[V0,dV]=analyticBS(m,Type1,Type2);

a=size(m);
f=(alp+(mm*(1-beta))^(-1))*(-Init_port-dV-alp);
% f=-dV-alp;
A=zeros(1,a(1,1));
b=0;
Aeq=[V0;dV]
beq=[1; ret]
LB=-100*ones(1,a(1,1));
UB=100*ones(1,a(1,1));
X0=zeros(20,1);
% interior-point, active-set, simplex
options=optimoptions(@linprog,'Algorithm', 'interior-point','MaxIter',100000)
[X,FVAL, EXITFLAG, OUTPUT,LAMDA]=linprog(f,A,b,Aeq,beq,LB,UB,X0,options);



% [X,FVAL, EXITFLAG, OUTPUT,LAMDA]=linprog(f,A,b,Aeq,beq,LB,UB);

risk=f*X;

%risk is calculated wihtout cost
%now we factor cost into min Prob

bar(X)
```

```
function z= zrisk(tdays, mu,m,Type1, Type2)
%Zrisk is used for the hedging problem of case study 2.

ret=mu;
alp=0.0;
beta=0.99;
mm=1;
w=0.005;


Init_port=Barrier_price(100,115,70,0.01,0.02,0.04,tdays,100,100);
[V0,dV]=analyticBS(m,Type1,Type2);

a=size(m);
f=(alp+(mm*(1-beta))^(-1))*(-Init_port-dV-alp);
% f=-dV-alp;
A=zeros(1,a(1,1));
b=0;
Aeq=[V0;dV]
beq=[1; ret]
LB=-100*ones(1,a(1,1));
UB=100*ones(1,a(1,1));
X0=zeros(20,1);
% interior-point, active-set, simplex
options=optimoptions(@linprog,'Algorithm', 'interior-point','MaxIter',100000);
[X,FVAL, EXITFLAG, OUTPUT,LAMDA]=linprog(f,A,b,Aeq,beq,LB,UB,X0,options);

EXITFLAG;

z=f*X;
end
```

```
%This code solves for the hedging problem. It investigates the time effects
%of barrier option in the risk and return characteristics.
%See the paper by Alexendra, Coleman, Li. This code
%corresponds to the Case Study 2 Part 2,
%Done By Resham Sivnarain, 11230712
clear X;
clear Y;
clear Z;
clear Yl;
clear Xl;

t=100;
i=1;
j=1;
k=0;
Yl=5:t;
% Xl=0.05:0.05:0.5;
Y=[];
X=[];
```

```
for r =0.05:0.05:0.5
    Xl=ones(1,t-4)*r;
    for td=5:t
       k=k+1;
       Z(k)=zrisk(td,r,m,Type1,Type2);

    end
    Y=[Y,Yl];
    X=[X,Xl];
    i=i+1;
end

X=reshape(X,t-4,10);
Y=reshape(Y,t-4,10);
Z=reshape(Z,t-4,10);

surf(X,Y,Z)
```

Case Study 3

```
%%
% this case study solves for Case study 3, CVaR optimisation with cost.
%Matlab function descriptions
% X = linprog(f,A,b,Aeq,beq,LB,UB) defines a set of lower and upper
%     bounds on the design variables, X, so that the solution is in
%     the range LB <= X <= UB. Use empty matrices for LB and UB
%     if no bounds exist. Set LB(i) = -Inf if X(i) is unbounded below;
%     set UB(i) = Inf if X(i) is unbounded above.

% X = linprog(f,A,b) attempts to solve the linear programming problem:
%
%              min f'*x    subject to:   A*x <= b
%               x
%
%     X = linprog(f,A,b,Aeq,beq) solves the problem above while additionally
%     satisfying the equality constraints Aeq*x = beq.


%first solve for a risk parameter with cost consideration
clear ret;
clear risk;
ret=0.56;
alp=0.0;
beta=0.99;
mm=20;
w=0.005;

[V0,dV]=analyticBS(m,Type1,Type2);

a=size(m);
f=(alp+(mm*(1-beta))^(-1))*(-dV-alp);
```

```
% f=-dV-alp;
A=zeros(1,a(1,1));
b=0;
Aeq=[V0;dV]
beq=[1; ret]
LB=-0.3*ones(1,a(1,1));
UB=0.8*ones(1,a(1,1));
X=linprog(f,A,b,Aeq,beq,LB,UB)
risk=f*X;

%risk is calculated wihtout cost
%now we factor cost into min Prob

% fcost=-dV-alp+w*abs(risk);
fcost=f+w*abs(risk);
A=zeros(1,a(1,1));
b=0;
Aeq=[V0;dV]
beq=[1; ret]
LB=-0.3*ones(1,a(1,1));
UB=0.8*ones(1,a(1,1));
X2=linprog(fcost,A,b,Aeq,beq,LB,UB);

bar(X)
figure
bar(X2)

% plot(risk,ret,'-o')

for alp=-10:20
    a(alp+11)=smoothing(alp,5)
    b(alp+11)=smoothing(alp,4)
end
plot(a)
hold on
plot (b,'r')
%%
hist(PL,80);
hold on
x = -5:0.1:5;
y = 50*tpdf(x,2);
z = normpdf(x,0,1);

plot(x,y,'-')
```

Case Study 4

```
data =FTSEJSEClose1;
roll=200;
%this program takes a data set and fits a distribution to the data subset
%which is rolling of 200 days of FTSE/JSE ALSI data. This code is used for
%the investigation into the stats effects on FTSE/JSE
%Done by Resham Sivnarain, 2016

aa=size(data);
data_end=aa(1,2)-roll-1;
% create a data series for drift, vol and degree of freedom
%also work out the CvaR of the data set
for i= 1: data_end
    d=data(i:i+roll);
    pd=fitdist(d,'tLocationScale')
    drft(i)=pd.mu;
    sigma(i)=pd.sigma;
    nu(i)=pd.nu;
    cvar(i)=my_cvar(d,5)

end
% a1=size(drft);

for i=1:(data_end-1)
   delta_drft(i) = (cvar(i+1)-cvar(i))/(drft(i+1)-drft(i));
   delta_sigma(i)=(cvar(i+1)-cvar(i))/( sigma(i+1)-sigma(i));
   delta_nu(i)=(cvar(i+1)-cvar(i))/(nu(i+1)-nu(i));

end

plot(delta_drft)
```

Case Study 5

```
%KNAPSACK Solves the 0-1 knapsack problem for positive integer weights
%
%  [BEST AMOUNT] = KNAPSACK(WEIGHTS, VALUES, CONSTRAINT)
%
%       WEIGHTS    : The weight of every item (1-by-N)
%       VALUES     : The value of every item (1-by-N)
%       CONSTRAINT : The weight constraint of the knapsack (scalar)
%
%       BEST       : Value of best possible knapsack (scalar)
%       AMOUNT     : 1-by-N vector specifying the amount to use of each item (0 or 1)
%
%
%   EXAMPLE :
%
%       weights = [1 1 1 1 2 2 3];
```

```
%       values  = [1 1 2 3 1 3 5];
%       [best amount] = KNAPSACK(weights, values, 7)
%
%       best =
%
%           13
%
%
%       amount =
%
%            0     0     1     1     0     1     1
%

function [best amount] = knapsack(weights, values, W)
    if ~all(is_positive_integer(weights)) || ...
       ~is_positive_integer(W)
        error('Weights must be positive integers');
    end
    %We work in one dimension
    [M N] = size(weights);
    weights = weights(:);
    values = values(:);
    if numel(weights) ~= numel(values)
        error('The size of weights must match the size of values');
    end
    if numel(W) > 1
        error('Only one constraint allowed');
    end

    % Solve the problem

    % Note that A would ideally be indexed from A(0..N,0..W) but MATLAB
    % does not allow this.
    A = zeros(length(weights)+1,W+1);
    % A(j+1,Y+1) means the value of the best knapsack with capacity Y using
    % the first j items.
    for j = 1:length(weights)
        for Y = 1:W
            if weights(j) > Y
                A(j+1,Y+1) = A(j,Y+1);
            else
                A(j+1,Y+1) = ...
                    max( A(j,Y+1), values(j) + A(j,Y-weights(j)+1));
            end
        end
    end

   best = A(end,end);

   %Now backtrack
   amount = zeros(length(weights),1);
   a = best;
```

```
   j = length(weights);
   Y = W;
   while a > 0
       while A(j+1,Y+1) == a
           j = j - 1;
       end
       j = j + 1; %This item has to be in the knapsack
       amount(j) = 1;
       Y = Y - weights(j);
       j = j - 1;
       a = A(j+1,Y+1);
   end



    amount = reshape(amount,M,N);
end

function yn = is_positive_integer(X)
    yn = X>0 & floor(X)==X;
end
```