COHERENT AND CONVEX MEASURES OF RISK

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF APPLIED MATHEMATICS

OF

THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

İREM YILDIRIM

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

THE DEPARTMENT OF FINANCIAL MATHEMATICS

SEPTEMBER 2005

Approval of the Graduate School of Applied Mathematics

Prof. Dr. Ersan AKYILDIZ
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. Hayri KÖREZLİOĞLU
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Prof. Dr. Hayri KÖREZLİOĞLU
Supervisor

Examining Committee Members

Prof. Dr. Hayri KÖREZLİOĞLU

Prof. Dr. Ömer GEBİZLİOĞLU

Assoc. Prof. Dr. Azize HAYFAVİ

Assist. Prof. Dr. Kasırga YILDIRAK

Dr. Coşkun KÜÇÜKÖZMEN

Abstract

COHERENT AND CONVEX MEASURES OF RISK

İrem Yıldırım
M.Sc., Department of Financial Mathematics
Supervisor: Prof. Dr. Hayri Körezlioğlu

September 2005, 77 pages

One of the financial risks an agent has to deal with is market risk. Market risk is caused by the uncertainty attached to asset values. There exit various measures trying to model market risk. The most widely accepted one is Value-at-Risk. However Value-at-Risk does not encourage portfolio diversification in general, whereas a consistent risk measure has to do so. In this work, risk measures satisfying these consistency conditions are examined within theoretical basis. Different types of coherent and convex risk measures are investigated. Moreover the extension of coherent risk measures to multiperiod settings is discussed.

Keywords: market risk, Value-at-Risk, coherent risk measures, convex risk measures, multi period dynamic model.

Öz

UYUMLU VE KONVEKS RİSK ÖLÇÜMLERİ

İrem Yıldırim
Yüksek Lisans, Finansal Matematik Bölümü
Tez Yöneticisi: Prof. Dr. Hayri Körezlioğlu

Eylül 2005, 77 sayfa

Piyasadaki oyuncuların maruz kaldığı finansal risklerden birisi de piyasa riskidir. Piyasa riski yatırım araçlarının gelecekte alabileceği değerlerin belirsizliğinden kaynaklanır. Bu riski belirlemeye çalışan pek çok model yapılmıştır. Bunların en yaygın olarak kullanılanı Riske Maruz Değerdir (RMD). Ancak tutarlı ölçülerin portföy çeşitlendirmesini desteklemesi beklenirken, RMD bunu yapmaz. Bu çalışmada bu tür tutarlılık koşullarını sağlayan risk ölçümlerinin teorik altyapısı ve uyumlu ve konveks risk ölçü çeşitleri incelenmiştir. Ayrıca uyumlu risk ölçülerinin çoklu döneme genişletilmesi araştırılmıştır.

Anahtar Kelimeler: piyasa riski, Riske Maruz Değer, uyumlu risk ölçüleri, konveks risk ölçüleri, ҫoklu dönem dinamik model.

To my family

ACKNOWLEDGMENTS

I am grateful to Prof. Dr. Hayri Körezlioğlu, Assoc. Prof. Dr. Azize Hayfavi, Assist. Prof. Dr. Kasırga Yıldırak and Dr. Coşkun Küçüközmen for patiently guiding, motivating, and encouraging me throughout this study.

I want to thank Zehra Ekşi without whom this thesis would not have been possible.

I am also indebted to Seval Çevik, Aslı Kunt, Yeliz Yolcu, Ayşegül İşcanoğlu, Serkan Zeytun, Uygar Pekerten and Oktay Sürücü.

Finally, I want to thank the members of the Institute of Applied Mathematics.

Table of Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

The most important question for an agent entering the market with an expectation of financial gain, is "how bad can it be?". Because in financial markets it is known that higher return is always related to higher risk. Therefore, the question arising first is what the market risk is. Unfortunately, defining and evaluating market risk is a hard work because risk is a qualitative element. However market risk can be defined as the degree of uncertainty of future net worths. When we use this definition, arising question is how can we measure this "degree of uncertainty". A market risk measure, trying to determine the degree of uncertainty, is the additional capital required to cover possible losses. According to this statement, we need to estimate the "possible losses" in order to measure the market risk of a financial position.

For decades many researchers have been trying to formalize an answer to measure the market risk. As a result there is a vast amount of literature in this field of study. Two most widely known tools used to formalize the market risk are Greeks, measuring the sensitivity of assets to market movements, and Value-at-Risk (VaR). Although Leavens did not unequivocally present VaR model, he can be regarded as the pioneer of early VaR studies. This is due to the fact that Leavens published the first and the most comprehensive study about the benefits of portfolio diversification in 1945. Markowitz(1952) and later Roy(1952) followed Leavens by publishing the same VaR measures independently. William

Sharpe proposed the Capital Asset Pricing Model in 1963. Thirty years after this, the committee formed by JP Morgan for a study on derivatives used the term Value-at-Risk firstly in a report published in 1993. In October 1994 JP Morgan proposed a new system called Risk Metrics. It was a free computer system providing risk measures for 400 financial instruments. Moreover following the approval of the limited use of VaR measures for calculating bank capital requirements in 1996 by the Basle Committee, VaR became the most widely used financial risk measure.

VaR is the maximum amount of loss that can be observed for a given confidence level in the determined time interval. For instance, if it is said that VaR of a position at 95% confidence level is 1000, this means that in 95 days out of 100 you can expect to face a loss lower than 1000. VaR is basically a quantile estimation for a determined probability distribution. For a continuous distribution with a given confidence level α

$$VaR_\alpha(X) = F^{-1}(\alpha)$$

,where F^{-1} is the inverse cumulative distribution function of losses of portfolio X. Also there is another formulation offered by Artzner, Deldean, Eber and Heath (ADEH). Since the original definition in [ADEH99] uses profit distribution, differently from the previous formulation we need a minus sign at the beginning. Moreover, this time since we have to work with the left of the graph, α should be taken as 0.05.

$$VaR_\alpha(X) = -inf(x|P(X \leq x) > \alpha)$$

The graphical representation gives a better insight. For instance in Figure 1.1 VaR at α confidence level is q_α^+. Unfortunately this definition of VaR does not encourage portfolio diversification. This means risk assigned to a composite portfolio can be higher than the sum of VaR numbers of the separate portfolios. An explanatory example can be found in [FS02b]. Such inconsistencies

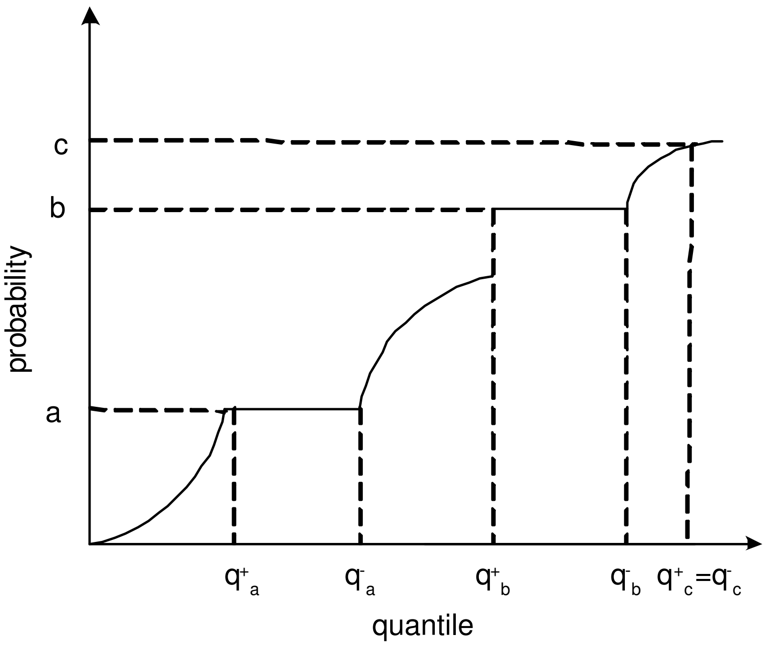

Figure 1.1: Value-at-Risk

of VaR motivated researchers to formalize better measures of risk. Some of them offered modifications and extensions in Value-at-Risk while others were offering alternative ways for financial risk calculation. The first line of research was started by Artzner, Deldean, Eber and Heath (ADEH) in 1997 with the article entitled "Thinking Coherently ". The major contribution of these scholars is the introduction of "Coherent Risk Measures" in 1999. These papers introduced the consistency conditions which should be satisfied by a sensible risk measure. Since VaR is not a coherent risk measure in the given context, new risk measures that both satisfy these consistency conditions and as easy to compute as VaR are constructed. Conditional Value at Risk (CVaR) by Uryasev and Rockafeller in 1999 and Expected Shortfall (ES) by Acerbi et. al. in 2000 are two examples. Both of these measures work with the α percent worst cases and take an expectation on these worst losses. After these, Artzner et. al. extended coherent risk measures to multiperiod setting in 2002. Another important contribution in this area was made by Föllmer and Schied in 2002 with the introduction of "Convex Risk Measures". These measures drop the positive homogeneity axiom of coherent risk measures, which assumes a linear relation between the size of a position and it's risk level. After these in 2004 Bion-Nadal introduced conditional convex risk measures which integrates the asymmetric information theory in risk measurement.

In Chapter 2 theoretical basis of coherent, convex and conditional convex risk measures will be given. Unfortunately having strong mathematical basis is not enough for these measures to compete with VaR. Due to this fact coherent and convex risk measures formalized for application purposes will be discussed in Chapter 3. Lastly the extension of coherent risk measures to multiperiod dynamic setting will be given in the Chapter 4.

Chapter 2

COHERENT, CONVEX AND CONDITIONAL CONVEX RISK MEASURES

Freddy Delbean, Chair of Financial Mathematics of ETH Zürih, says in one of his speeches that making a definition of risk is extremely complicated if not impossible, but a theory is needed to make decisions to deal with the availability of money under uncertainty. He continues by saying people told them it was not possible to represent financial risk by just one number. However a yes or no answer is needed under uncertainty. As an answer coherent risk measures were formulated.

Definition 2.1. A measure of risk is a mapping from $\mathcal{X}$ into $\mathbb{R}$ i.e.

$$\rho : \mathcal{X} \to \mathbb{R}$$

The risk measures that will be investigated in this chapter are the results of an ambitious work to reach more compatible results with the financial environment. As said in the previous chapter, no matter how advanced the techniques used in its computation are, VaR suffers from not being a convex mapping (unless asset returns are assumed to be normally distributed). It is obvious that a risk measure, which does not encourage portfolio diversification is in contradiction with the entire related financial literature starting from the portfolio theory of Markowitz.

Deficiencies of VaR motivated ADEH to construct a new model said to satisfy some consistency conditions determined by the authors. In the first section of this chapter the line of reasoning for this model will be given. Afterwards Föllmer and Schied formed convex measures of risk, which also reflect the liquidity risk in the financial markets. Section two will introduce these models. A relatively new contribution was done by Bion-Nadal. Convex risk measures were evolved by the addition of asymmetric information theory in financial markets. These conditional convex risk measures will constitute the last section of this chapter.

2.1 Coherent Measures of Risk

When an investor, a manager or a supervision agency has to decide whether to take a specific position in the market or not, he evaluates possible portfolio returns that can be faced under different scenarios. This is done to decide whether the position is acceptable or not. Therefore there must be a boundary or a minimum return level requirement for a position to be acceptable. After defining the boundary condition, portfolios satisfying this condition will be said to compose the acceptance set. When a position is labelled as unacceptable, an investor can totally give this option up or by adding some risk free asset to the portfolio, he can make it acceptable. The cost of acquiring this risk free asset is used to measure the risk of a given portfolio. The aim of this chapter is to explain the risk measure satisfying the consistency conditions defined by ADEH and answering the question of how much additional capital is needed to make the position acceptable. This will be done in two parts; firstly, the set of possible states of world at the end of the period will assumed to be finite and secondly, it will be assumed that it is infinite. Throughout the paper the risk free rate of return is assumed to be zero, so risk free investment is the cash added to the portfolio.

2.1.1 Coherent Risk Measures when Ω is Finite

In this subsection it will be assumed that all possible states of world are known. However the probabilities of occurrence for these states are not. The set of possible states of world at the end of the period is denoted by Ω and it is assumed to have a finite number of elements. $\mathcal{X}$ represents all real valued functions on Ω. If card(Ω)= n, $\mathcal{X}$ can be identified with $\mathbb{R}^n$.

Definition 2.2. A risk measure ρ is a coherent risk measure if it satisfies

1. Monotonicity: For all X and Y $\in \mathcal{X}$; if $X \leq Y$, $\rho(X) \geq \rho(Y)$.
2. Translation Invariance: For all X $\in \mathcal{X}$ and for all real numbers α;

   $\rho(X + \alpha) = \rho(X) - \alpha$.
3. Positive Homogeneity: For all $\lambda \geq 0$ and for all $X \in \mathcal{X}$; $\rho(\lambda X) = \lambda \rho(X)$
4. Subadditivity: For all $X_1, X_2 \in \mathcal{X}$; $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$.

Monotonicity implies that if a portfolio has higher returns for all possible states of nature relative to another portfolio, risk associated to this portfolio is naturally lower. Translation invariance assures that a risk measure is expressed in currency terms by saying that when α amount of cash is added to the portfolio as a risk free investment, the risk of the portfolio decreases by the same amount. Positive homogeneity implies that there is a linear relation between the position size and the associated risk of the portfolio. Lastly subadditivity says that a merger does not create extra risk.

Definition 2.3. Acceptance set associated to a risk measure ρ is

$$\mathcal{A}_\rho = \{X \in \mathcal{X} \mid \rho(X) \leq 0\}.$$

Definition 2.4. Coherent risk measure associated to an acceptance set is denoted by

$$\rho_{\mathcal{A}}(X) = inf\{m \mid (m + X) \in \mathcal{A}\}.$$

Proposition 2.1. If a set $\mathcal{B}$ satisfies the following conditions,

1. $\mathcal{B}$ contains the cone of non-negative elements in $\mathcal{X}$, L_+,
2. $\mathcal{B}$ does not intersect with the set L_{--}, where

$$\mathrm{L}_{--} = \{X \mid \ for\ each\ \omega \in \Omega, X(\omega) < 0\},$$

3. $\mathcal{B}$ is convex,
4. $\mathcal{B}$ is a positively homogenous cone,

associated risk measure, $\rho_{\mathcal{B}}$ is coherent. Moreover $\mathcal{A}_{\rho_{\mathcal{B}}} = \bar{\mathcal{B}}$.

Proof: 1) When $\parallel \ \parallel$ represents supremum norm, $- \parallel X \parallel \leq X \leq \parallel X \parallel$ for all $X \in \mathcal{X}$. Since $\parallel X \parallel + X \geq 0$, there is a real number m>$\parallel X \parallel$. Then $m + X \in \mathrm{L}_+$. Therefore $\rho(X) \leq \parallel X \parallel$. On the other hand $X - \parallel X \parallel \leq 0$. If $m < - \parallel X \parallel$, then $m + X \notin \mathrm{L}_+$ and $\rho(X) \geq \parallel X \parallel$. As a result $\rho_{\mathcal{B}}(X)$ is a finite number.
2) There exists real numbers p, q such that $inf\{p \mid X + (\alpha + p) \in \mathbb{B}\} = inf\{q \mid X + q \in \mathcal{B}\} - \alpha$ for $\forall \alpha \in \mathbb{R}$. Therefore $\rho_{\mathcal{B}}(X)$ satisfies translation invariance.
3) For $X, Y \in \mathcal{X}$; but not in $\mathcal{B}$, by property 3 if $X + m$, $Y + n \in \mathcal{B}$, then $\alpha(X + m) + (Y + n) \in \mathcal{B}$, where $\alpha \in [0, 1]$. Take $\alpha = 1/2$, by property 4 if $1/2(X + Y + m + n) \in \mathcal{B}$ then $(X + Y + m + n) \in \mathcal{B}$. Since $\{s : (X + Y) + s \in \mathcal{B}\} \supset \{m : X + m \in \mathcal{B}\} + \{n : Y + n \in \mathcal{B}\}$, $\rho_{\mathcal{B}}(X + Y) \leq \rho_{\mathcal{B}}(X) + \rho_{\mathcal{B}}(Y)$. This means $\rho_{\mathcal{B}}$ satisfies subadditivity.
4) For a real number m; if $m > \rho_{\mathcal{B}}(X)$, then for each $\lambda > 0$, $\lambda X + \lambda m \in \mathcal{B}$. Therefore $\rho_{\mathcal{B}}(\lambda X) \leq \lambda m$. If $m < \rho_{\mathcal{B}}(X)$, then for each $\lambda > 0$, $\lambda X + \lambda m \notin \mathcal{B}$ and $\rho_{\mathcal{B}}(\lambda X) \geq \lambda m$. Therefore $\rho_{\mathcal{B}}(\lambda X) = \lambda \rho_{\mathcal{B}}(X)$. As a result $\rho_{\mathcal{B}}$ satisfies positive

homogeneity.
5) If $X(\omega) \leq Y(\omega)$ for all $\omega \in \Omega$ and $X + m \in \mathcal{B}$, then also $Y + m \in \mathcal{B}$. $Y + m = X + m + (Y - X)$. Since $Y - X \geq 0$, by the property 1 $Y - X \in \mathcal{B}$. $\{m : m + X \in \mathcal{B}\} \subset \{m : m + Y \in \mathcal{B}\}$ therefore $\rho_{\mathcal{B}}(X) \geq \rho_{\mathcal{B}}(Y)$ and satisfies monotonicity.
6) For each $X \in \mathcal{B}$, $\rho_{\mathcal{B}}(X) \leq 0$ since $X \in \mathrm{L}_+$ and $\rho_{\mathcal{B}}$ satisfies monotonicity and positive homogeneity. Proposition below assures that $\mathcal{A}_{\rho_{\mathcal{B}}}$ is closed, which proves that $\mathcal{A}_{\rho_{\mathcal{B}}} = \bar{\mathcal{B}}$.

Proposition 2.2. If ρ is a coherent risk measure, then $\mathcal{A}_\rho$ is closed and satisfies properties 1-4 of *Proposition 1.1.* Moreover $\rho = \rho_{\mathcal{A}_\rho}$.

Proof: 1) From subadditivity and positive homogeneity, $\rho(\alpha X + (1-\alpha)Y) \leq \alpha\rho(X) + (1-\alpha)\rho(Y)$ for $\alpha \in [0,1]$. This implies that ρ is a convex function. Hence it is continuous. Therefore $\mathcal{A}_\rho = \{X \mid \rho(X) \leq 0\}$ is closed. If $X, Y \in \mathcal{A}_\rho$; i.e. $\rho(X) \leq 0$, $\rho(Y) \leq 0$, $\rho(\alpha X + (1-\alpha)Y) \leq 0$. Therefore $\alpha X + (1-\alpha)Y \in \mathcal{A}_\rho$. If $\rho(X) \leq 0$ and $\lambda > 0$, then $\rho(\lambda X) = \lambda\rho(X) \leq 0$ and $\lambda X \in \mathcal{A}_\rho$. Therefore $\mathcal{A}_\rho$ is a closed convex positively homogenous cone.
2) From positive homogeneity $\rho(0) = 0$. If $X(\omega) \geq 0$ $\forall\omega \in \Omega$, then $\rho(X) \leq 0$ by monotonicity. Therefore $\mathrm{L}_+ \supset \mathcal{A}_\rho$.
3) Let $X \in \mathrm{L}_{--}$ and $\rho(X) < 0$. However monotonicity of ρ implies that $\rho(X) \geq 0$ since $X < 0$; this is a contradiction. If $\rho(X) = 0$ for α such that $\alpha > 0$ and $X + \alpha \in \mathrm{L}_{--}$, then $\rho(X + \alpha) = \rho(X) - \alpha < 0$. This is also a contradiction. Hence $\rho(X) > 0$, $X \notin \mathcal{A}_\rho$. This means $\mathrm{L}_{--} \cap \mathcal{A}_\rho = \emptyset$.
4) $\rho_{\mathcal{A}_\rho}(X) = inf\{m \in \mathbb{R} \mid m + X \in \mathcal{A}_\rho\}$. In other words $\rho_{\mathcal{A}_\rho}(X) = inf\{m \in \mathbb{R} \mid \rho(X) \leq m\}$. Therefore $\rho_{\mathcal{A}_\rho}(X) = \rho(X)$.

In order to see the relation between coherent risk measures and acceptance sets, the illustration from [K04] is given in Figure 2.1. In this illustration it is assumed that Ω has only two elements. That is $\Omega = \{\omega_1, \omega_2\}$. Two axes represent different values of $X_1(\omega)$ and $X_2(\omega)$. $X_1 = X_1(\omega_1)$, $X_2 = X_2(\omega)$. Two cases are given in the figure; when $X \in \mathcal{A}$ and when $X \notin \mathcal{A}$.

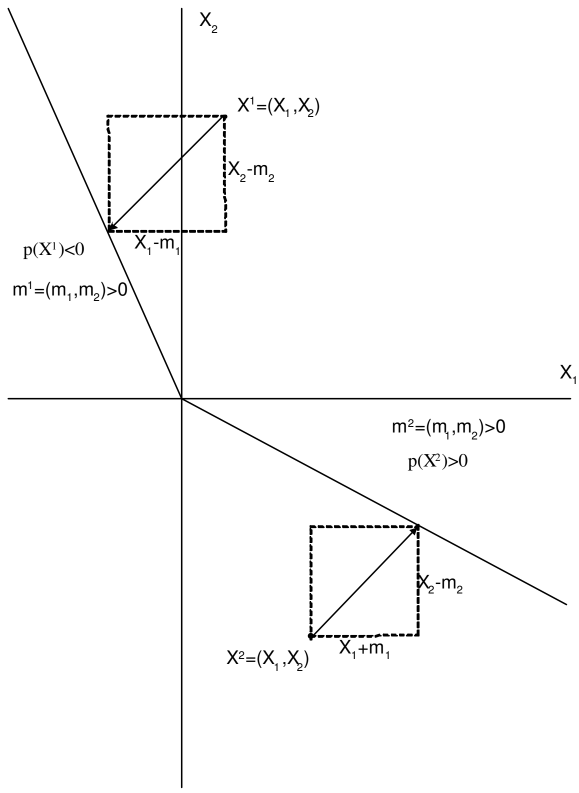

Figure 2.1: Coherent Measures of Risk

Remark: Convexity of an acceptance set means that the portfolio at the global minimum can be found. This property is important for the optimization used for portfolio selection. Not having a convex acceptance set, VaR cannot assure a global minimum.

According to [ADEH99] any coherent risk measure arises as the supremum of expected negative returns (losses) for some collection of generalized scenarios or probability measures on Ω.

Proposition 2.3. A risk measure is coherent if and only if there exists a family $\mathcal{P}$ of probability measures on Ω such that:

$$\rho(X) = sup\{E_P[-X] \mid P \in \mathcal{P}\}.$$

Proof: First part of the theorem is obvious. If $\rho(X) = sup\{E_P[-X] \mid P \in \mathcal{P}\}$, then:

i) If $X(\omega) \leq Y(\omega)$ for $\forall \omega \in \Omega$, then $E_P[-X] \geq E_P[-Y]$ for $\forall P \in \mathcal{P}$. This implies monotonicity.

ii) For any constant α; $E_P[-(X+\alpha)] = E_P[-X] - \alpha$ for $\forall P \in \mathcal{P}$. Taking the supremum preserves the inequality and translation invariance of $\rho(X)$ follows.

iii) For a real number $\lambda > 0$; $E_P[-(\lambda X)] = \lambda E_P[-X]$ for $\forall P \in \mathcal{P}$. This implies positive homogeneity.

iv) For $X, Y \in \mathcal{X}$; $sup\{E[-(X+Y)] \mid P \in \mathcal{P}\} \leq sup\{E[-X] \mid P \in \mathcal{P}\} + sup\{E[-Y] \mid P \in \mathcal{P}\}$. This gives subadditivity.

Conversely, let $\mathcal{M}$ denotes the set of all probability measures on Ω. Define $\mathcal{P}_\rho$ as

$$\mathcal{P}_\rho = \{P \in \mathcal{M} : \forall X \in \mathcal{X}, E[-X] \leq \rho(X)\}$$

,where ρ is assumed to be a coherent risk measure. The set of probabilities $\mathcal{M}$ is a compact set in $\mathbb{R}^n$, where $n = card(\Omega)$, since it is a closed subset of unit ball in $\mathbb{R}^n$, which is compact. In fact $\mathcal{M} = \{P \in \mathbb{R}^n : \forall \omega; P(\omega) \geq 0$ and $\sum P(\omega) = 1\}$. Given $X \in \mathcal{X}$, $E[-X]$ is continuous from $\mathcal{M}$ into $\mathbb{R}$, due to the fact that a continuous image of a compact set $\{E[-X] : P \in \mathcal{M}\}$ is compact in $\mathbb{R}$. This implies $\{E[-X] : P \in \mathcal{M}\} \cap \{a \in \mathbb{R} : a \leq \rho(X)\}$ is compact since a closed subset of a compact metric space is compact. Therefore

$$\rho(X) = sup\{E_P[-X] \mid P \in \mathcal{P}_\rho\}.$$

2.1.2 Coherent Risk Measures when Ω is Infinite

Assuming that possible states of nature are finite is not compatible with daily financial market conditions. In reality nothing is impossible, even a plain crashing into the Twin Towers of New York. Observing a market movement equal to the one caused by this accident is nearly zero in statistical terms. Therefore restricting possible states of the world only to a finite list of asset prices cannot give a helpful sight of the risk associated to the position held. When the size of a position is millions of dollars, no one wants to think that there can be possibilities that are skipped. Due to these facts, extending coherent risk measures to infinite Ω is a necessity.

This part is based on [D00] that aims to extend the notion of coherent risk measures into arbitrary probability spaces. The finite dimensional space $\mathbb{R}^\Omega$ representing $\mathcal{X}$ is replaced with the space of all bounded measurable functions $\mathrm{L}^\infty(\Omega, \mathcal{F}, P)$, where P is an a priori given probability measure. This P does not mean every agent has a common view on the distribution of portfolio returns, but there exists a common view about the null sets.

Other notations that will be used throughout this section are: $\mathrm{L}^1(\Omega, \mathcal{F}, P)$ representing all integrable real random variables, $\mathrm{L}^\infty = (\mathrm{L}^1)'$ meaning L^∞ is the dual of L^1 (Appendix, A2). $(\mathrm{L}^\infty)'$ = Banach space $\mathbf{ba}(\Omega, \mathcal{F}, P)$ of all bounded finitely additive measures (Appendix, A1) $\mathcal{M}$ on $(\Omega, \mathcal{F})$ which are absolutely continuous with respect to P.

Definition 2.5. A mapping $\rho : \mathrm{L}^\infty(\Omega, \mathcal{F}, P) \to \mathbb{R}$ is called a coherent measure of risk if it satisfies the following conditions:

1. If $X \geq 0$, then $\rho(X) \leq 0$.
2. Subadditivity: For all $X_1, X_2 \in \mathcal{X}$; $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$.
3. Positive Homogeneity: For all $X \in \mathcal{X}$ and $\lambda \geq 0$; $\rho(\lambda + X) = \lambda\rho(X)$.
4. For all $X \in \mathcal{X}$ and every constant function a; $\rho(X + a) = \rho(X) - a$.

Theorem 2.1. Suppose that $\rho : \mathrm{L}^\infty(\Omega, \mathcal{F}, P) \to \mathbb{R}$ is a coherent measure of risk. There is a convex $\sigma(\mathbf{ba}(\Omega, \mathcal{F}, P), \mathrm{L}^\infty(\Omega, \mathcal{F}, P))$ (Appendix, A3) closed set $\mathcal{P}_{ba}$ of finitely additive probabilities such that:

$$\rho(X) = \sup_{\mathcal{M} \in \mathcal{P}_{ba}} E_{\mathcal{M}}[-X]$$

,where $\mathbf{ba}(\Omega, \mathcal{F}, P)$ refers to all bounded, finitely additive measures $\mathcal{M}$ on $(\Omega, \mathcal{F})$ which are absolutely continuous with respect to P. $\mathcal{M} \in \mathbf{ba}(\Omega, \mathcal{F}, P)$ is a finitely additive probability measure if $\mathcal{M}(1) = 1$.

Proof: 1) $\mathcal{C} = \{X \mid \rho(X) \leq 0\}$ is clearly a norm (supremum norm) closed positively homogenous cone in L^∞. Moreover ${\mathrm{L}^\infty}_+ \subset \mathcal{C}$.
2) The polar set $\mathcal{C}^o = \{\mathcal{M} : \forall X \in \mathcal{C}; E[X] \geq 0\}$ (Appendix, A4) is also a convex cone closed for *weak** topology on $\mathbf{ba}(\Omega, \mathcal{F}, P)$. (Appendix, Remark) Indeed let $\mathcal{M}_1, \mathcal{M}_2 \in \mathcal{C}^o$ and $E_{\mathcal{M}_1}[X] \geq 0, E_{\mathcal{M}_2}[X] \geq 0$. For $0 \leq \alpha \leq 1$

$$E_{\alpha\mathcal{M}_1+(1-\alpha)\mathcal{M}_2}[X] = \alpha \int_\Omega X d\mathcal{M}_1 + (1-\alpha) \int_\Omega X d\mathcal{M}_2$$

Since this expression is non negative, $\mathcal{C}^o$ is convex. For $\lambda \geq 0$ and $\int_\Omega X d\mathcal{M} \geq 0$; $\int X d(\lambda\mathcal{M}) = \lambda \int X d\mathcal{M}$. Therefore if $\mathcal{M} \in \mathcal{C}^o$, then $\lambda\mathcal{M} \in \mathcal{C}^o$. Moreover by the Proposition 9 of [RR73](Appendix, A4), $\mathcal{C}^o$ is absolutely continuous and $\sigma(\mathbf{ba}(\Omega, \mathcal{F}, P), \mathrm{L}^\infty(\Omega, \mathcal{F}, P))$ closed.
3) All elements in $\mathcal{C}^o$ are positive since $\mathrm{L}^\infty_+ \subset \mathcal{C}$. By the Proposition 9 of [RR73], $\mathcal{C}^o \subset (\mathrm{L}^\infty_+)^o \subset (\mathrm{L}^\infty_+)' = \mathbf{ba}(\Omega, \mathcal{F}, P)$. By definition $(\mathrm{L}^\infty_+)^o = \{\mathcal{M} \mid \forall X \in \mathrm{L}^\infty_+, E_{\mathcal{M}} \geq 0\}$. Therefore $\mathcal{C}^o$ contains only positive measures. This implies that for the set $\mathcal{P}_{ba} = \{\mathcal{M} \mid \mathcal{M} \in \mathcal{C}^o$ and $\mathcal{M}(1) = 1\}$ since $\mathcal{C}^o$ is a positive cone $\mathcal{C}^o = \bigcup_{\lambda \geq 0} \lambda\mathcal{P}_{ba}$.
4) According to the bipolar theorem in [Sch73] (Appendix, A5); $\mathcal{C} = \{X \mid \forall\mathcal{M} \in \mathcal{P}_{ba}, E_{\mathcal{M}}[X] \geq 0\}$. Indeed by the Robertson polar theorem $(\mathcal{C}^o)^o = \bigcap \mathcal{P}^o_{ba}$. $\mathcal{C}$ is a convex set containing 0, by the bipolar theorem $\mathcal{C} = \mathcal{P}^o_{ba} = \{X \mid \forall\mathcal{M} \in \mathcal{P}_{ba}, E_{\mathcal{M}}[X] \geq 0\}$.
5) The steps above imply that $\rho(X) \leq 0$ if and only if $E_{\mathcal{M}} \geq 0$ for all $\mathcal{M} \in \mathcal{P}_{ba}$. $X + \rho(X) \in \mathcal{A}_\rho$. Therefore for $\forall\mathcal{M} \in \mathcal{P}_{ba}$, $E_{\mathcal{M}}[X + \rho(X)] \geq 0$. As a result $\sup_{\mathcal{M}\in\mathcal{P}_{ba}} E_{\mathcal{M}}[-X] \leq \rho(X)$.
6) For an arbitrary $\varepsilon > 0$, $\rho(X + \rho(X) - \varepsilon) > 0$ and $X + \rho(X) - \varepsilon \notin \mathcal{A}_\rho$. Therefore there is an $\mathcal{M} \in \mathcal{P}_{ba}$ such that $E_{\mathcal{M}}[X + \rho(X) - \varepsilon] < 0$, which leads to opposite equality.

As a result

$$\rho(X) = \sup_{\mathcal{M}\in\mathcal{P}_{ba}} E_{\mathcal{M}}[-X].$$

The relation between $\mathcal{C}$ and ρ is given by $\rho(X) = \inf\{\alpha \mid X + \alpha \in \mathcal{C}\}$.
In the theorem above the representation of the coherent risk measures is given in terms of finitely additive probability measures. In order to extend this to σ-finite probability measures extra conditions are needed.

Definition 2.6. A risk measure ρ is said to satisfy Fatou property if $\rho(X) \leq \liminf \rho(X_n)$ for any sequence, $(X_n)_{n\geq 1}$, of functions uniformly bounded by 1 and converging to X in probability.

Theorem 2.2. For a coherent risk measure ρ, the following are equivalent:

1. There is an $\mathrm{L}^1(\Omega, \mathcal{F}, P)$ closed convex set of probability measures $\mathcal{P}_\sigma$, all being absolutely continuous with respect to P and such that for $X \in \mathrm{L}^\infty$:

$$\rho(X) = \sup_{\mathcal{Q}\in\mathcal{P}_\sigma} E_{\mathcal{Q}}[-X].$$

2. The convex cone $\mathcal{C} = \{X \mid \rho(X) \leq 0\}$ is $\sigma(\mathrm{L}^\infty(P), \mathrm{L}^1(P))$ closed.

3. ρ satisfies the Fatou property.

Proof: 2⇒3 If $\mathcal{C}$ is $\sigma(\mathrm{L}^\infty(P), \mathrm{L}^1(P))$ closed, then ρ satisfies the Fatou property. If (X_n) is an increasing sequence converging to X in probability and $\| X_n \| \leq 1$ (where $\| \ \|$ is supremum norm) for all n and $\rho(X_n)$ decreases to some limit a, then $X_n + \rho(X_n) \in \mathcal{C}$. Since $\mathcal{C}$ is $\sigma(\mathrm{L}^\infty(P), \mathrm{L}^1(P))$ closed, limit $X + a$ is in $\mathcal{C}$. This implies $\rho(X + a) \leq 0$, so $\rho(X) \leq a$.

3⇒2 If ρ satisfies the Fatou property, then $\mathcal{C}$ is $\sigma(\mathrm{L}^\infty(P), \mathrm{L}^1(P))$ closed.
According to [G73]; if $E = \mathrm{L}^1(\mu)$, where μ is finitely countable measure, and H a convex subset of dual $E' = \mathrm{L}^\infty(\mu)$, H is weakly closed (i.e. H is closed for bounded sequences converging almost everywhere). Let $(X_n)_n \geq 1$ be a

sequence in $\mathcal{C}$, bounded by 1 and converging to X in probability. Since $\mathcal{C}$ is a convex cone, it is sufficient to check $\mathcal{C} \cap \mathcal{B}_1$ is closed in probability, where $\mathcal{B}_1$ is the closed unit ball in L^∞. By the Fatou property $\rho(X) \leq liminf \rho(X_n) \leq 0$. Therefore X is also in $\mathcal{C}$.

2⇒1 This part is parallel to the proof of Theorem 1.1.
Let $\mathcal{C} = \{X \mid \rho(X) \leq 0\}$. It is a norm closed cone in L^∞. Moreover $\mathrm{L}^\infty \subset \mathcal{C}$. By the second property $\mathcal{C}$ is $\sigma(\mathrm{L}^\infty(P), \mathrm{L}^1(P))$ closed.

The polar set of $\mathcal{C}$, $\mathcal{C}^o = \{f \mid f \in \mathrm{L}^1 \textit{ and } E_P[fX] \geq 0 \textit{ for all } X \in \mathcal{C}\}$ is also a convex cone closed for $\sigma(\mathrm{L}^1(P), \mathrm{L}^\infty(P))$. Indeed, for $f_1, f_2 \in \mathcal{C}^o$ and $E_P[f_1X \geq 0]$ and $E_P[f_2X \geq 0]$; if $0 \leq \alpha \leq 1$, $E_P[\alpha f_1 X + (1-\alpha) f_2 X] = \alpha E_P[f_1X] + (1-\alpha)E_P[f_2X] \geq 0$. Therefore $\mathcal{C}^o$ is a convex set. For $\lambda \geq 0$ and $E_P[fX] \geq 0$; $E_P[\lambda f X] = \lambda E_P[fX] \geq 0$. Therefore $\mathcal{C}^o$ is a positively homogeneous cone. Again by [RR73], $\mathcal{C}^o$ is absolutely convex and $\sigma(\mathrm{L}^1(P), \mathrm{L}^\infty(P))$ closed.

Since P is a probability measure and $\mathrm{L}^\infty_+ \subset \mathcal{C}$, $\mathcal{C}^o$ has only positive elements. For the set $\mathcal{P}_\sigma$ which is defined as $\mathcal{P}_\sigma = \{f \mid dQ = dPf \textit{ defines a prob-ability measure and } f \in \mathcal{C}^o\}$.

Since $\mathcal{C}^o$ is a positively homogenous cone, $\mathcal{C}^o = \bigcup_{\lambda \geq 0} \lambda P_\sigma$. By the bipolar theorem; $\mathcal{C} = \mathcal{P}_\sigma^o = \{X \mid \forall f \in \mathcal{P}_\sigma : E_P[fX] \geq 0\}$. Equivalently, $\mathcal{C} = \mathcal{P}_\sigma^o = \{X \mid \forall Q \in \mathcal{P}_\sigma : E_Q[X] \geq 0\}$. This implies that $\rho(X) \leq 0$ if and only if $E_Q[X] \geq 0$ for all $Q \in \mathcal{P}_\sigma$. Given that for every X; $X + \rho(X) \in \mathcal{A}_\rho$. Then $E_Q[X + \rho(X)] \geq 0$ and $\sup_Q E_Q[-X] \leq \rho(X)$.

For an arbitrary $\varepsilon > 0$, $\rho(X + \rho(X) - \varepsilon) > 0$ and $X + \rho(X) - \varepsilon \notin \mathcal{A}_\rho$. Therefore there is an $\mathcal{M} \in \mathcal{P}_\sigma$ such that $E_\mathcal{M}[X + \rho(X) - \varepsilon] < 0$, which leads to the opposite equality.

1⇒2 According to the Fatou lemma in [KHa01], let X_n be a sequence of extended real random variables and X be an integrable extended real random variable. If $X_n \leq X$ equivalently $-X_n \geq -X$, then for each $Q \in \mathcal{P}_\sigma$ $liminf\ E[-X_n] \geq E[\lim inf(-X_n)] \geq E[-X]$. When X_n is taken as a sequence uniformly bounded by 1 and tending to X in probability,the above inequality gives the Fatou property.

Remark: All of the possible positions used in this section are assumed to be bounded. This means that although there are some possible states of the world in Ω considering extreme scenarios, our positions do not take $+/-\infty$ values. In order to consider such a situation instead of L^∞, L^0 should be used. L^0 is the space of all real valued random variables. When X is a very risky position, there will not be enough capital to make this portfolio acceptable. In this case ρ will be $+\infty$. However when L^0 is used, situations that can lead to $\rho = -\infty$ should be avoided. Since there is no positions still being acceptable, after an infinite amount of capital is drawn out.

2.2 Convex Measures of Risk

The coherent risk measures of ADEH were a milestone in quantifying the risk of a position. Although there are many other measures for quantifying risk, consistency conditions brought by the coherent risk measures are so widely accepted that most of the risk measures are evaluated with respect to these conditions. The convex risk measures proposed by Föllmer *et. al.* is an example. It forms a theoretical framework going one step further in terms of reflecting real market conditions. Main idea is that the positive homogeneity axiom of coherent risk measures assumes that the risk of a financial position is linearly related to its size. For instance, if the size of a portfolio is doubled, the associated risk is twice the risk of the original portfolio. Föllmer *et. al.* argue that such an assumption ignores liquidity risk in financial markets. If the market cannot assure liquidity, the risk exposure of the portfolio grows faster than its volume. When the liquidity risk is taken as 0 in convex risk measures, the resulting measure will be coherent. Therefore, coherency is a special case of convex risk measures.

Ω represents the set of possible states of the world. A financial position denoted as $X : \Omega \to \mathbb{R}$. X belongs to the class $\mathcal{X}$ of financial positions, where $\mathcal{X}$ is taken as the linear space of bounded functions on Ω. $\mathcal{X}$ is assumed to include all constants and closed under the addition of constants. There is no probability measure given a priori.

Definition 2.7. Mapping $\rho : \mathcal{X} \to \mathbb{R}$ is a convex risk measure if it satisfies

1. Monotonicity: For all $X, Y \in \mathcal{X}$; if $X \leq Y$, $\rho(X) \geq \rho(Y)$.
2. Translation Invariance: For all $X \in \mathcal{X}$; if $m \in \mathbb{R}$, then $\rho(Y + m) = \rho(Y) - m$.
3. Convexity: For all $X, Y \in \mathcal{X}$; $\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y)$ for any $\lambda \in [0, 1]$.

Convexity property says that diversification decreases the risk of the portfolio.

Remark: When a risk measure satisfies positive homogeneity, convexity implies subadditivity. Take $\lambda = 1/2$,

$$\rho(1/2(X+Y)) \leq \rho(1/2X) + \rho(1/2Y).$$

By positive homogeneity

$$1/2\rho(X+Y) \leq 1/2(\rho(X) + \rho(Y)).$$

Remark: Convex risk measure $\rho(X)$ is said to be normalized if $\rho(0) = 0$. Then $\rho(X)$ is the amount of money invested in risk free asset to make position X acceptable.

Any risk measure $\rho : \mathcal{X} \to \mathbb{R}$ implies its own acceptance set.

$$\mathcal{A}_\rho = \{X \in \mathcal{X} \mid \rho(X) \leq 0\}.$$

For any class of acceptable positions $\mathcal{A}$, a convex set measure ρ can be defined as:

$$\rho(X) = \inf\{m \mid m + X \in \mathcal{A}\}.$$

The following propositions summarizing the relation between acceptance set and risk measure, are parallel to the ones given for coherent risk measures.

Proposition 2.4. Suppose $\rho : \mathcal{X} \to \mathbb{R}$ is a convex risk measure with associated acceptance set $\mathcal{A}_\rho$. Then $\rho_{\mathcal{A}_\rho} = \rho$. Moreover,

1. $\mathcal{A}_\rho$ is a non empty set satisfying

$$\inf\{m \in \mathbb{R} \mid m \in \mathcal{A}\} \geq -\infty. \qquad (1.1)$$

2. Hederitary Property: If $X \in \mathcal{A}_\rho$ and $Y \in \mathcal{X}$, satisfying $Y \geq X$, then $Y \in \mathcal{A}_\rho$.

3. If $X \in \mathcal{A}$ and $Y \in \mathcal{X}$, then $\{\lambda \in [0,1] \mid |\lambda X + (1-\lambda)Y \in \mathcal{A}\}$ is closed in [0,1].

4. ρ is a convex measure of risk if and only if $\mathcal{A}$ is convex.

Proof: The first part of the proof is similar to the fourth step in proof of Proposition 1.2. For the properties of $\mathcal{A}_\rho$;
1) $0 \in \mathcal{X}$, $\rho(0 + \rho(0)) = \rho(0) - 0 \leq 0$. Therefore, $0 \in \mathcal{A}_\rho$. Since $X : \Omega \to \mathbb{R}$ for $\forall X \in \mathcal{X}$; if $\rho(X) = -\infty$, $\rho(X - \infty) \leq 0$. This means that although an infinite amount of money is drawn from the portfolio it is still riskless. Such a portfolio does not exist.
2) If $X \in \mathcal{A}_\rho$, $\rho(X) \leq 0$. If $Y \geq X$ for all $\omega \in \Omega$ by monotonicity $\rho(Y) \leq \rho(X) \leq 0$, then $Y \in \mathcal{A}_\rho$.

3) The function $\lambda \to \rho(\lambda X+(1-\lambda)Y)$ is continuous since it is convex and takes only finite values. Hence set of $\lambda \in [0,1]$, such that $\rho(\lambda X + (1-\lambda)Y) \leq 0$, is closed.

4) For $X, Y \in \mathcal{A}$, by the convexity of ρ, $\rho(\lambda X + (1-\lambda)Y) \leq 0$. Hence $\lambda X + (1-\lambda)Y \in \mathcal{A}$. This means $\mathcal{A}$ is a convex set. The converse part of the property follows Proposition 1.5 given below.

Proposition 2.5. Assume $\mathcal{A} \neq \emptyset$ is a convex subset of $\mathcal{X}$, satisfying hederitary property and inequality (1.1). If $\rho_{\mathcal{A}} = \inf\{m \in \mathbb{R} \mid m + X \in \mathcal{A}\}$, then

1. $\rho_{\mathcal{A}}$ is a convex measure of risk.
2. $\mathcal{A}$ is a subset of $\mathcal{A}_{\rho_{\mathcal{A}}}$. Moreover if $\mathcal{A}$ satisfies the third property of Proposition 1.4, then $\mathcal{A} = \mathcal{A}_{\rho_{\mathcal{A}}}$.

Proof: 1) If $X, Y \in \mathcal{X}$, then $X + \rho_{\mathcal{A}}(X)$ and $Y + \rho_{\mathcal{A}}(Y) \in \mathcal{A}$. From the convexity of $\mathcal{A}$, $\lambda(X + \rho_{\mathcal{A}}(X)) + (1-\lambda)(Y + \rho_{\mathcal{A}}(Y)) \in \mathcal{A}$. Therefore $\lambda\rho_{\mathcal{A}}(X) + (1-\lambda)\rho_{\mathcal{A}}(Y) \geq \rho_{\mathcal{A}}(\lambda X + (1-\lambda)Y)$ gives the convexity of $\rho_{\mathcal{A}}$. For portfolio X; $\rho_{\mathcal{A}}(X) = \inf\{m \in \mathbb{R} \mid m + X \in \mathcal{A}\}$. If positive k amount of cash is added to the portfolio, $\rho_{\mathcal{A}}(X+k) = \inf\{m - k \in \mathbb{R} \mid m - k + (X+k) \in \mathcal{A}\}$. Hence translation invariance follows. If $Y \geq X$ for all $\omega \in \Omega$, then $\rho_{\mathcal{A}}(X) = \inf\{m \in \mathbb{R} \mid m + X \in \mathcal{A}\} \geq \inf\{m \in \mathbb{R} \mid m + Y \in \mathcal{A}\} = \rho_{\mathcal{A}}(Y)$, means that $\rho_{\mathcal{A}}$ satisfies the monotonicity condition. To show that $\rho_{\mathcal{A}}$ takes only finite values, fix some Y of non empty set $\mathcal{A}$. For every $X \in \mathcal{X}$, there exists a finite number m with $m+X > Y$ since X and Y are bounded. By monotonicity $\rho(m+X) \leq \rho(Y) \leq 0$. By translation invariance $\rho(X) \leq m$. To show $\rho_{\mathcal{A}}(X) > -\infty$, take m' such that $X + m' \leq 0$. $\rho(X + m') \geq \rho(0) \; - \infty$. Therefore $\rho(X) > -\infty$.

2) $\mathcal{A}_{\rho_{\mathcal{A}}} = \{X \mid \inf\{m \in \mathbb{R} \mid m + X \in \mathcal{A}\} \leq 0\} \supseteq \mathcal{A}$. Let $\mathcal{A}$ satisfy the third property of Proposition 1.5. It must be shown that if $X \notin \mathcal{A}$, then $\rho_{\mathcal{A}} > 0$. Take $m > \rho_{\mathcal{A}}(0)$, $\rho_{\mathcal{A}}(0) - m < 0$, $\rho_{\mathcal{A}}(m) < 0$. There exists an $\varepsilon \in [0,1]$ such that $\varepsilon m + (1-\varepsilon)X \notin \mathcal{A}$. Thus $\rho_{\mathcal{A}}(\varepsilon m + (1-\varepsilon)X) \geq 0$, $\rho_{\mathcal{A}}((1-\varepsilon)X) - \varepsilon m \geq 0$, $\varepsilon m \leq \rho_{\mathcal{A}}(1-\varepsilon)X) = \rho_{\mathcal{A}}(\varepsilon 0 + (1-\varepsilon)X) \leq \rho_{\mathcal{A}}(0) + (1-\varepsilon)\rho_{\mathcal{A}}(0)$. As a result $\rho_{\mathcal{A}}(X) \geq \frac{\varepsilon(m - \rho_{\mathcal{A}}(0))}{1-\varepsilon} > 0$.

2.2.1 Convex Risk Measures when Ω is Finite

Theorem 2.3. Suppose $\mathcal{X}$ is the space of all real valued functions on a finite set Ω. Then $\rho : \mathcal{X} \to \mathbb{R}$ is a convex measure of risk if and only if there is a penalty function $\alpha : \mathcal{P} \to (-\infty, \infty]$ such that

$$\rho(Z) = \sup_{Q \in \mathcal{P}} (E_Q[-Z] - \alpha(Q)).$$

The function α satisfies $\alpha(Q) \geq \rho(0)$ for any $Q \in \mathcal{P}$ for any $Q \in \mathcal{P}$.

Proof: The if part: For each $Q \in \mathcal{P}$; the risk function is defined as

$$\rho : X \to E_Q(-X) - \alpha(Q).$$

This expression is: 1) *Convex*: $\lambda X + (1-\lambda)Y \to E_Q[-(\lambda X + (1-\lambda)Y)] - \alpha(Q) \leq \lambda E_Q[-X - \alpha(Q)] + (1-\lambda)E_Q[-Y - \alpha(Q)]$. This inequality is preserved when the supremum is taken.

2) *Translation invariant*: $X + k \to E_Q[-(X + k) - \alpha(Q)] = E_Q[-X] - k - \alpha(Q)$.

3) *Monotone*: If $X \leq Y$, $E_Q[-X] - \alpha(Q) \geq E_Q[-Y] - \alpha(Q)$. Taking supremum preserves the inequality.

For the converse implication, define $\alpha(Q)$ for $Q \in \mathcal{P}$ as

$$\alpha(Q) = \sup_{X \in \mathcal{X}} (E_Q[-X] - \rho(X)).$$

Then claim that $\alpha(Q) = \sup_{X \in \mathcal{A}_\rho} E_Q[-X]$. Since $\mathcal{A}_\rho \subseteq \mathcal{X}$, $\sup_{X \in \mathcal{X}}(E_Q[-X] - \rho(X)) \geq \sup_{X \in \mathcal{A}_\rho} E_Q[-X]$. To show the converse inequality take an arbitrary X and say $X' = X + \rho(X) \in \mathcal{A}_\rho$. Then

$$\sup_{X \in \mathcal{A}_\rho} E_Q[-X] \geq E_Q[-X'] = E_Q[-X] - \rho(X),$$

$$\sup_{X \in \mathcal{A}_\rho} E_Q[-X] \geq \sup_{X \in \mathcal{X}} (E_Q[-X] - \rho(X)).$$

This proves the claim. Note that the assumption of Ω being finite has not been used yet.

Fix some $Y \in \mathcal{X}$ and take α as defined above. Then

$$E_Q[-Y] - \rho(Y) \leq \sup_{X \in \mathcal{X}} (E_Q[-X] - \rho(X)),$$

$$\rho(Y) \geq E_Q[-Y] - \alpha(Q),$$

$$\rho(Y) \geq \sup_{Q \in \mathcal{P}} (E_Q[-Y] - \alpha(Q)).$$

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

$$m > \sup_{Q \in \mathcal{P}} (E_Q(-Y) - \alpha(Q)).$$

It must be shown that $m \geq \rho(Y)$ or equivalently $m + Y \in \mathcal{A}_\rho$. This will be done by contradiction. Suppose that $m + Y \notin \mathcal{A}_\rho$. By definition being convex function on $\mathbb{R}^n$, where $n = card(\Omega)$, ρ takes only finite values. Moreover according to Rockafeller's Convex Analysis [1], ρ is continuous. Here $\rho(X) \leq 0$ forms a convex set $\mathcal{A}_\rho$. By the separation theorem (Appendix, A6)a linear functional can be found such that

$$\beta : \sup_{X \in \mathcal{A}_\rho} l(X) < l(m + Y) =: \gamma < \infty$$

claiming that l follows a negative linear functional. Indeed normalization and monotonicity axiom imply that ; $\rho(0) = 0$ and $\rho(X) \leq \rho(0)$ for $X \geq 0$. Thus if $X \in \mathcal{X}$ satisfies $X \geq 0$, then $\lambda X + \rho(0) \notin \mathcal{A}_\rho$ for all $\lambda \geq 1$. Hence $\gamma > l(\lambda X + \rho(0)) = \lambda l(X) + \rho(0)$. As $\lambda \nearrow \infty$ inequality holds only if $l(X) \leq 0$.

Without loss of generality it can be assumed that $l(1) = -1$. Then $Q(A) := l(-\mathbf{I}_A)$ defines a probability measure $Q \in \mathcal{P}$. Furthermore $l(X) = E_Q[-X]$ for

[1]From [Ro70]; Corollary 10.1.1: A convex function finite on all of $\mathbb{R}^n$ is necessarily continuous.

$\forall X \in \mathcal{A}_\rho$. Taking the supremum of both sides preserves equality;

$$\sup_{X\in\mathcal{A}_\rho} l(X) = \sup_{X\in\mathcal{A}_\rho} E_Q(-X).$$

By definition $\beta = \alpha(Q)$. But $E_Q[-Y] - m = l(m+Y) = \gamma > \beta = \alpha(Q)$. This is a contradiction, so $m \geq \rho(Y)$ and $m + Y \in \mathcal{A}_\rho$.

2.2.2 Convex Risk Measures when Ω is Infinite

When Ω is infinite, $\mathcal{X}$, the space of possible financial positions, is taken to be the linear space of all bounded measurable functions on a measurable space $(\Omega, \mathcal{F})$. The line of reasoning is as follow; firstly since integration is well defined on a finitely additive, non-negative set function on the given linear space ([DuSc58], III.2 of 6), the representation theorem will be constructed on these. Next the theory will be extended in order to pass to the σ finite (Appendix,A1) probability measures. Such a representation can be examined under two different assumptions; there may either be an a priori probability measure on $(\Omega, \mathcal{F})$ which means that $\mathcal{X} = \mathrm{L}^\infty(\Omega, \mathcal{F}, P)$, in other words between the agents there is at least a common view on which events are not likely to occur. Or there may be complete uncertainty, no probability measure in advance given.

Necessary notations for this part are:

- $\mathcal{M}_{1,f} = \mathcal{M}_{1,f}(\Omega, \mathcal{F})$: The class of all finitely additive probability measures on $\mathcal{F}$ that are normalized to 1 i.e. $Q(\Omega) = 1$.
- $\mathcal{M}_1 = \mathcal{M}_1(\Omega, \mathcal{F})$: The class of all probability measures on $(\Omega, \mathcal{F})$.
- $\alpha : \mathcal{M}_{1,f} \to \mathbb{R} \cup \{\infty\}$: The penalty function which is not identically equal to ∞.

Definition 2.8. When there is no a priori given probability measure, for each

$Q \in \mathcal{M}_{1,f}$; ρ is defined as follows

$$\rho = \sup_{Q \in \mathcal{M}_{1,f}} (E_Q[-X] - \alpha(Q)) \qquad (1.2).$$

When defined like this, ρ is

1. *Monotone*: If $X \leq Y$, then for each Q; $E_Q[-X]-\alpha(Q) \geq E_Q[-Y]-\alpha(Q)$. Taking the supremum preserves inequality, then $\rho(X) \geq \rho(Y)$.

2. *Translation invariant*: $m \in \mathbb{R}$, $E_Q[-(X+m)] - \alpha(Q) = E_Q[-X] - m - \alpha(Q)$. When the supremum is taken, $\rho(X+m) = \rho(X) - m$.

3. *Convex*: $\rho(\lambda X + (1-\lambda)Y) = \sup_{Q \in \mathcal{M}_{1,f}} (E_Q[(-\lambda X + -(1-\lambda)Y)] - \alpha(Q))$

   $= \sup_{Q \in \mathcal{M}_{1,f}} (\lambda(E_Q[-X] - \alpha(Q)) + (1-\lambda)(E_Q[-Y] - \alpha(Q)))$
   $\leq \lambda \sup_{Q \in \mathcal{M}_{1,f}} (E_Q[-X] - \alpha(Q)) + (1-\lambda) \sup_{Q \in \mathcal{M}_{1,f}} (E_Q[-Y] - \alpha(Q))$
   $= \lambda\rho(X) + (1-\lambda)\rho(Y)$

It must be stated that there is no unique α. A penalty function characterizes the risk measure ρ it belongs. Therefore it is said that ρ is represented by α on $\mathcal{M}_{1,f}$.

Theorem 2.4. For any convex risk measure defined as (1.2) the penalty function, α_{min}, given as

$$\alpha_{min}(Q) = \sup_{X \in \mathcal{A}_\rho} E_Q[-X] \qquad for\ all\ Q \in \mathcal{M}_{1,f}$$

is the minimal penalty function representing ρ. That is any penalty function α for which (1.2) holds, $\alpha(Q) \geq \alpha_{min}(Q)$ is satisfied for all $Q \in \mathcal{M}_{1,f}$.

Proof: From the representation given in (1.2), for all $Q \in \mathcal{M}_{1,f}$; $\rho(X) \geq E_Q[-X] - \alpha(Q)$. Therefore, $\alpha(Q) \geq \sup_{X \in \mathcal{A}_\rho}(E_Q[-X] - \rho(X))$. If $X \in \mathcal{A}_\rho$, $\rho(X) \leq 0$. Hence $\alpha(Q) \geq \sup_{X \in \mathcal{A}_\rho} E_Q[-X]$. As a result, $\alpha(Q) \geq \alpha_{min}(Q)$ for all $Q \in \mathcal{M}_{1,f}$.

Remark: Coherent risk measures can be given as a special case of convex risk measures represented by the penalty function α, when α is defined as:

$$\begin{cases} 0 \text{ if } \mathrm{Q} \in \mathcal{Q} \\ +\infty \text{ otherwise} \end{cases}$$

,where $\mathcal{Q} \subset \mathcal{M}_{(1,f)}$. Hence $\sup_{Q\in\mathcal{Q}}(E_Q[-X] - \alpha(Q)) = \sup_{Q\in\mathcal{Q}} E_Q[-X]$.

The next step in [FS02c] is representing a convex risk measure using σ-finite probability measures instead of finitely additive non-negative set functions. This is done by constructing a penalty function α taking infinite values when $Q \notin \mathcal{M}_1(\Omega, \mathcal{F})$. Hence, $\rho(X) = \sup_{Q\in\mathcal{M}_{(1,f)}}(E_Q[-X] - \alpha(Q))$, where $\alpha(Q) = \infty$ if $Q \notin \mathcal{M}_1(\Omega, \mathcal{F})$. As a result

$$\rho(X) = \sup_{Q\in\mathcal{M}_1} (E_Q[-X] - \alpha(Q)). \qquad (1.3)$$

This kind of representation is closely related with certain continuity properties given in [FS02b] as follows.

Lemma 2.1. A convex risk measure ρ, represented as (1.3) is continuous from above in the sense that

$$X_n \searrow X \Longrightarrow \rho(X_n) \nearrow \rho(X). \qquad (1.4)$$

This condition is also equivalent to; if X_n is a bounded sequence in $\mathcal{X}$ converging pointwise to $X \in \mathcal{X}$, then

$$\rho(X) \leq \liminf_{n\nearrow\infty} \rho(X_n). \qquad (1.5)$$

Proof: Firstly (1.5) holds if ρ can be represented in terms of probability measures. The dominated convergence theorem says that, if $\{X_n, n \in N\}$ converges if and only if there is an integrable extended random variable X such that $\forall n$

$\mid X_n \mid \leq X$, then

$$\begin{aligned} E[\lim_{n\to\infty} X_n] &= \lim_{n\to\infty} E[X_n], \\ E[X] &= \lim_{n\to\infty} E[X_n], \\ E[X_n] &\to E[X] \; as \; n \to \infty. \end{aligned}$$

This implies for each $Q \in \mathcal{M}_1$ $E_Q[X_n] \to E_Q[X]$ as $n \to \infty$. Hence

$$\begin{aligned} \rho(X) &= \sup_{Q\in\mathcal{M}_1} (\lim_{n\to\infty} E_Q[-X_n - \alpha(Q)]), \\ E_Q[-X_n - \alpha(Q)] &\leq \sup_{Q\in\mathcal{M}_1} (\lim_{n\to\infty} E_Q[-X_n - \alpha(Q)]), \\ \lim_{n\to\infty} (E_Q[-X_n - \alpha(Q)]) &\leq \liminf_{n\to\infty} \sup_{Q\in\mathcal{M}_1} (E_Q[-X_n - \alpha(Q)]), \\ \sup_{Q\in\mathcal{M}_1} (\lim_{n\to\infty} (E_Q[-X_n] - \alpha(Q))) &\leq \liminf_{n\to\infty} \sup (E_Q[-X_n] - \alpha(Q)), \\ &= \liminf_{n\to\infty} \rho(X_n). \end{aligned}$$

To show the equivalence of (1.4) and (1.5), we will firstly assume that (1.5) holds. For each n; if $X_n \searrow X$, then $\rho(X) \geq \rho(X_n)$. As a result $\rho(X_n) \nearrow \rho(X)$.

If (1.4) holds (i.e. (X_n) is a bounded sequence in $\mathcal{X}$, converging to X), then define $Y_m = \sup_{n\geq m} X_n \in \mathcal{X}$. This implies Y_m decreases P a.s. to X. Since $\rho(X_n) \geq \rho(Y_n)$, by monotonicity, (1.4) yields that

$$\liminf_{n\nearrow\infty} \rho(X_n) \geq \lim_{n\nearrow\infty} \rho(Y_n) = \rho(X).$$

The above lemma says that if ρ is concentrated on probability measures, it satisfies continuity from above. A stronger condition, continuity from below, says that if all increasing sequences in $\mathcal{X}$ are continuous from below, then "any penalty function" is concentrated on $\mathcal{M}_1$.

Proposition 2.6. Let ρ be a convex measure of risk which is continuous from below in the sense that

$$X_n \nearrow X \Rightarrow \rho(X_n) \searrow \rho(X)$$

and α a penalty function on $\mathcal{M}_1$, representing ρ. Then,

$$\alpha(Q) < \infty \Rightarrow Q \ is \ \sigma \ additive.$$

Proof: The proof this proposition can be found in [FS02b] pg 169.

If there exists a specified probabilistic model on $(\Omega, \mathcal{F})$, $\mathcal{X}$ becomes $\mathrm{L}^\infty(\Omega, \mathcal{F}, P)$. Then it can be considered that

$$\rho(X) = \rho(Y) \text{ if } X = Y \text{ P a.s.} \qquad (1.6)$$

Lemma 2.2. Let ρ be a convex measure of risk satisfying (1.6) and represented by $\alpha(Q)$. Then for any probability measure which is not absolutely continuous with respect to P, $\alpha(Q) = \infty$.

Proof: If $Q \in \mathcal{M}_1(\Omega, \mathcal{F})$ is not absolutely continuous with respect to P, then there exists an $A \in \mathcal{F}$ such that $Q(A) > 0$, where $P(A) = 0$. Take any $X \in \mathcal{A}_\rho$ and define $X_n = X - nI_A$. Then $\rho(X_n) = \rho(X)$ since $X - nI_A = X$ P a.s. Therefore $X_n \in \mathcal{A}_\rho$ for $n \in N$.

$$\alpha(Q) \geq \alpha_{min}(Q) \geq E_Q[-X_n] = E_Q[-X] + nQ(A) \rightarrow \infty \ as \ n \rightarrow \infty$$

This means that if probability measures absolutely continuous with respect to P are denoted by $\mathcal{M}_1(P)$, ρ of any $X \in \mathrm{L}^\infty$ can be represented by penalty functions restricted to $\mathcal{M}_1(P)$.

$$\rho(X) = \sup_{Q \in \mathcal{M}_1(P)} (E_Q[-X] - \alpha(Q))$$

2.3 Convex Conditional Risk Measures

In each section of this chapter an assumption that is not compatible with real market conditions are dropped. Conditional convex risk measures integrates the asymmetric information theory, which has an important place in financial literature, to the theory of measuring financial market risk. There are various kinds of agents in the market: managers, low scale investors, supervisors, banks, funds and so on. Assuming that all these have access to the same information set is not realistic. Differentiation of the information sets that each investor has access to is called asymmetric information. When trying to model risk this difference should be considered. The theoretical structure of this integration was created by Bion-Nadal in her article dated June 2004. This section is based on this paper, [BN04].

The integration is carried out by using the conditional expectation, where the condition represents all accessible information. As it was already defined, $\mathcal{X}$ represents the set of financial positions, (which is) the linear space of bounded maps on Ω. But from now on the investor does not have full information on the entire σ-algebra of Ω. As a result he does not have access to all maps defined on Ω, but only to measurable maps defined on σ algebra $\mathcal{F}$ (which is not equal to $\sigma(\Omega)$). And the risk measure is formed as the conditional expectation of $X \in \mathcal{X}$ given the σ algebra $\mathcal{F}$.

Different from the previous sections, in this section risk measures will not be investigated when Ω consists of finite number of scenarios. Instead of this, restriction is done in terms of σ-algebra by assuming the entire σ-algebra is not known. Representation theorems for conditional convex risk measures will be given in two subsections; firstly under the assumption that there is complete uncertainty (i.e. no given probability measures); secondly assuming partial uncertainty, there exists an a priori given probability measure on the measurable space $(\Omega, \mathcal{F})$.

2.3.1 Convex Conditional Risk Measures Under Complete Uncertainty

The necessary building blocks for this subsection are the linear space $\mathcal{X}$ of financial positions where a financial position is a bounded map defined on Ω and sub-sigma algebra $\mathcal{F}$ on Ω. Then $(\Omega, \mathcal{F})$ consists a measurable space and $\mathcal{E}_{\mathcal{F}}$ denotes the set of all bounded real valued measurable maps on $(\Omega, \mathcal{F})$. There is no common view on which sets of $\mathcal{F}$ are null sets. In other words, there is no base probability given.

Definition 2.9. A mapping $\rho_{\mathcal{F}} : \mathcal{X} \longrightarrow \mathcal{E}_{\mathcal{F}}$ is called a risk measure conditional to the σ-algebra $\mathcal{F}$ if it satisfies the following:

1. *Monotonicity*: For all $X, Y \in \mathcal{X}$; if $X \leq Y$, then $\rho_{\mathcal{F}}(Y) \leq \rho_{\mathcal{F}}(X)$.
2. *Translation Invariance*: For all $X \in \mathcal{X}$ and $Y \in \mathcal{E}_{\mathcal{F}}$; $\rho(X+Y) = \rho(X) - Y$.
3. *Multiplicative Invariance*: For all $X \in \mathcal{X}$ and for all $A \in \mathcal{F}$; $\rho(XI_A) = I_A \rho_{\mathcal{F}}(X)$.

Remark: Contrary to the previous sections, in conditional risk measures, the risk of a position is not described with a single number but with a $(\Omega, \mathcal{F})$ measurable map. The risk measure of a position X conditional to σ algebra $\mathcal{F}$ is the minimal $\mathcal{F}$ measurable map, which, added to the initial position X, makes the position acceptable. This is a different from adding a determined level of cash which brings the same risk free return under each scenario. It can be said that conditional convex risk measures prevent the investor from holding an idle amount of capital to make the position acceptable.

Remark: The interpretation of first two properties , monotonicity and translation invariance, are the same as in the previous sections. Multiplicative invariance assures that if a position is acceptable when the whole σ algebra $\mathcal{F}$ is considered, then it is acceptable through each and every subset of $\mathcal{F}$.

Definition 2.10. A risk measure conditional to σ algebra $\mathcal{F}$ is a convex risk measure conditional to σ algebra $\mathcal{F}$ if it satisfies

- Convexity: For all $X, Y \in \mathcal{X}$ and $\lambda \in [0,1]$; $\rho_{\mathcal{F}}(\lambda X + (1-\lambda)Y) \leq \lambda\rho_{\mathcal{F}}(X) + (1-\lambda)\rho_{\mathcal{F}}(Y)$.

Definition 2.11. The acceptance set for a conditional risk measure is defined as

$$\mathcal{A}_{\rho_{\mathcal{F}}} = \{X \in \mathcal{X} \mid \rho_{\mathcal{F}}(X) \leq 0\}.$$

Proposition 2.7. The acceptance set $\mathcal{A} = \mathcal{A}_{\rho_{\mathcal{F}}}$ of a convex conditional risk measure $\rho_{\mathcal{F}}$

1. is non-empty, closed with respect to the supremum norm and has a hederitary property: for all $X \in \mathcal{A}$ and for all $Y \in \mathcal{X}$ if $Y \geq X$, then $Y \in \mathcal{A}$
2. satisfies the bifurcation property: for all $X_1, X_2 \in \mathcal{A}$ and for all disjoint $B_1, B_2 \in \mathcal{F}$

   $$X = X_1 I_{B_1} + X_2 I_{B_2}$$

   is in $\mathcal{A}$.
3. Every $\mathcal{F}$ measurable element of $\mathcal{A}$ is positive.
4. $\rho_{\mathcal{F}}$ can be recovered from $\mathcal{A}$

   $$\rho_{\mathcal{F}}(X) = \inf\{Y \in \mathcal{E}_{\mathcal{F}} \mid X + Y \in \mathcal{A}\}.$$

Proof: For the proof of this proposition [BN04], pg 7.

For the rest of this section it will be assumed that there is a σ algebra $\mathcal{G}$ such that $\mathcal{X}$ is the set of all bounded measurable functions on the measurable space $(\Omega, \mathcal{G})$ and $\mathcal{F}$ is a sub-sigma algebra of $\mathcal{G}$. $\mathcal{M}_{1,f}$ denotes the set of all finitely additive set functions. $Q : \mathcal{G} \to [0,1]$ such that $Q(\Omega) = 1$.

Theorem 2.5. Let $\rho_{\mathcal{F}}$ be a convex risk measure conditional to $\mathcal{F}$. Then, for all $X \in \mathcal{X}$ there is $Q_X \in \mathcal{M}_{1,f}$, such that for all $B \in \mathcal{F}$

$$E_{Q_X}[\rho_{\mathcal{F}}(XI_B)] = E_{Q_X}[-XI_B] - \sup_{Y \in \mathcal{A}_{\rho_{\mathcal{F}}}} E_Q[-YI_B]. \qquad (1.7)$$

For all $X \in \mathcal{X}$, for all $Q \in \mathcal{M}_{1,f}$ and for all $B \in \mathcal{F}$

$$E_Q[\rho_{\mathcal{F}}(XI_B)] \geq E_Q[-XI_B] - \sup_{Y \in \mathcal{A}_{\rho_{\mathcal{F}}}} E_Q[-YI_B]. \qquad (1.8)$$

Proof: For all $X \in \mathcal{X}$, $\rho_{\mathcal{F}}(\rho_{\mathcal{F}}(X) + X) = \rho_{\mathcal{F}}(X) - \rho_{\mathcal{F}}(X) = 0$. So for $\rho_{\mathcal{F}}(X) + X \in \mathcal{A}_{\rho_{\mathcal{F}}}$. Then for all $Q \in \mathcal{M}_{1,f}$

$$\sup_{Y \in \mathcal{A}_{\rho_{\mathcal{F}}}} E_Q[-Y] \geq E_Q[-X - \rho_{\mathcal{F}}(X)].$$

Putting a condition preserves the inequality for all $B \in \mathcal{F}$.

$$\sup_{Y \in \mathcal{A}_{\rho_{\mathcal{F}}}} E_Q[-YI_B] \geq E_Q[-XI_B] - E_Q[\rho_{\mathcal{F}}(X)I_B].$$

This gives us the inequality (1.8). It is enough to prove the equality (1.7) for $\rho_{\mathcal{F}}(X) = 0$. If $\rho_{\mathcal{F}}(X) \neq 0$, by replacing X by $X + \rho_{\mathcal{F}}(X)$, the same result can be reached. To begin, consider the convex hull C of $\{(Y - X)I_B, \rho_{\mathcal{F}}(Y) \leq 0 \text{ } and \text{ } B \in \mathcal{F}\}$ (i.e. the smallest convex set containing $\{(Y - X)I_B, \rho_{\mathcal{F}}(Y) \leq 0 \text{ } and \text{ } B \in \mathcal{F}\}$).

STEP 1: Prove that $C \cap \{0\} = \emptyset$

Assume that (there are) $\lambda_i \geq 0$; $\Sigma_{i=1}^n \lambda_i = 1$ and $\Sigma_{i=1}^n \lambda_i (Y_i - X) I_{B_i} = 0$. Choose $J \subset \{1, 2, ..., n\}$ such that $\tilde{B} = \cap_{i \in J} B_i \neq \emptyset$ and such that $\forall j \in \{1, 2, ..., n\} - J$, $\tilde{B} = \cap B_j = \emptyset$.

It is given that $\lambda_1(Y_1 - X)I_{B_1} + ... + \lambda_n(Y_n - X)I_{B_n} = 0$. This is the definition of the convex hull and, moreover, $I_{\cap_{i\in J}B_i}$ is not identically 0. Therefore

$$\lambda_1(Y_1-X)I_{B_1-\tilde{B}}+\lambda_1(Y_1-X)I_{\tilde{B}}+...+\lambda_n(Y_n-X)I_{B_n-\tilde{B}}+\lambda_n(Y_n-X)I_{\tilde{B}}=0$$

$$\Sigma_{i=1}^n\lambda_i(Y_i-X)I_{B_i-\tilde{B}}+\Sigma_{i=1}^n\lambda_i(Y_i-X)I_{\tilde{B}}=0$$

$$\Sigma_{i=1}^n\lambda_i(Y_i-X)I_{B_i-\tilde{B}}+\Sigma_{i\in J}\lambda_i(Y_i-X)I_{\tilde{B}}=0$$

The condition that $\rho_{\mathcal{F}}(Y)\leq 0$ and $\rho_{\mathcal{F}}(X)=0$ gives that for all $Y\geq X$ and $\lambda_i\geq 0$, where $i=1,...,n$, the above equation is satisfied if both of the summations are equal to 0.

Let $\tilde{Y}=\frac{\Sigma_{i\in J}\lambda_iY_i}{\Sigma_{i\in J}\lambda_i}=\frac{\lambda 1Y_1}{\Sigma_{i\in J}\lambda_i}+...+\frac{\lambda nY_n}{\Sigma_{i\in J}\lambda_i}$ If $\frac{\lambda i}{\Sigma_{i\in J}}=k_i$, then $\Sigma_{i\in J}k_i=1$. From the convexity of $\rho_{\mathcal{F}}$, $\rho_{\mathcal{F}}(\tilde{Y})<0$. On the other hand $\tilde{Y}I_{\tilde{B}}=XI_{\tilde{B}}$ so $\rho_{\mathcal{F}}(\tilde{Y}I_{\tilde{B}})=\rho_{\mathcal{F}}(XI_{\tilde{B}})=0$. This is a contradiction. Therefore the convex hull C does not contain $\{0\}$.

STEP 2: C contains the open ball $B_1(1-X)=\{Y\in\mathcal{X};\parallel Y-(1-X)\parallel<1\}$. Indeed if $Y\in B_1(1-X)$, $\parallel Y-X-1\parallel<1$. Call $Y+X=Z$, then $\parallel Z-1\parallel<1$. We can define $B_1(1)$ as follows; $B_1(1)=\{Z\in\mathcal{X};\parallel Z-1\parallel<1\}$. Because of the supremum norm given, $-1<Z-1<1$. By monotonicity $\rho(2)<\rho(Z)<\rho(0)$. Since $\rho(0)=0$, $\rho(Z)<0$. Recalling $Y=Z-X$ the claim follows.

STEP 3: Prove the existence of $Q_X\in\mathcal{M}_{1,f}$ such that $\forall B\in\mathcal{F}$ $E_{Q_X}[-XI_B]=\sup_{Y\in\mathcal{A}_{\rho_{\mathcal{F}}}}E_{Q_X}[-YI_B]$.
From step 1, 0 does not belong to C and C is a non-empty set. As a consequence of the Separation Theorem in Appendix A6, there exists a non-zero continuous linear functional l on $\mathcal{X}$, such that $0=l(0)\leq l(Z)$ for all $Z\in C$. $0\leq l((Y-X)I_B)$ for all Y satisfying $\rho_{\mathcal{F}}(Y)<0$ and for all $B\in\mathcal{F}$. For all $Y\in\mathcal{A}_{\rho_{\mathcal{F}}}$ and $\forall\epsilon>0$; $\rho(Y+\epsilon)<0$. Hence by the continuity of l,

$$0\leq l((Y-X)I_B)\ \forall Y\in\mathcal{A}_{\rho_{\mathcal{F}}}. \qquad (1.9)$$

Now for $\forall Y \geq 0$ and $\forall \lambda > 0$, $\rho_{\mathcal{F}}(1 + \lambda Y) < 0$. Therefore $(1 + \lambda Y - X) \in C$ and $l(1 + \lambda Y - X) = l(1) + \lambda l(Y) - l(X) \geq 0$. To verify this inequality for every $\lambda > 0$, $l(Y)$ must be greater than or equal to 0. This means that l is a positive functional. Therefore $l(1) > 0$. From Appendix A7, there is a unique $Q_X \in \mathcal{M}_{1,f}$ defined as $E_{Q_X}(Y) = l(Y)/l(1)$ for all $Y \in \mathcal{X}$. From (1.9), $E_{Q_X}(-XI_B) \geq E_{Q_X}(-YI_B)$ for all $B \in \mathcal{F}$ and $\forall Y \in \mathcal{A}_{\rho_{\mathcal{F}}}$. This implies

$$E_{Q_X}[-XI_B] \geq \sup_{Y \in \mathcal{A}_{\rho_{\mathcal{F}}}} E_{Q_X}[-YI_B].$$

From the inequality (1.8), $E_{Q_X}[-XI_B] - \sup_{Y \in \mathcal{A}_{\rho_{\mathcal{F}}}} E_{Q_X}[-YI_B] \leq E_{Q_X}[\rho_{\mathcal{F}}(X)I_B]$. Since $\rho_{\mathcal{F}}(X) = 0$, $E_{Q_X}[\rho_{\mathcal{F}}(X)I_B] = 0$. This provides

$$E_{Q_X}[-XI_B] \leq \sup_{Y \in \mathcal{A}_{\rho_{\mathcal{F}}}} E_{Q_X}[-YI_B].$$

This ends step 3 and the proof.

Like the previous section, next step is making the necessary transformations to define $\rho_{\mathcal{F}}$ in terms of σ finite probability measures.

Definition 2.12. A convex conditional risk measure is continuous from below if for all increasing sequence X_n of elements of $\mathcal{X}$ converging to X, the decreasing sequence $\rho_{\mathcal{F}}(X_n)$ converges to $\rho_{\mathcal{F}}(X)$.

Theorem 2.6. Let $\rho_{\mathcal{F}}$ be a convex risk measure conditional to $\mathcal{F}$. Assume that $\rho_{\mathcal{F}}$ is continuous from below. Then for all $X \in \mathcal{X}$, for every probability measure P on $(\Omega, \mathcal{F})$; there is a Q_X in $\mathcal{M}_1(\mathcal{G}, \mathcal{F}, P)$ such that

$$\rho_{\mathcal{F}}(X) = E_{Q_X}[-X \mid \mathcal{F}] - \alpha(Q_X) \ P \ a.s.$$

,where $\mathcal{M}_1(\mathcal{G}, \mathcal{F}, P)$ is the set of all probability measure Q on $(\Omega, \mathcal{G})$ such that the restriction of Q to $\mathcal{F}$ is equal to P.

Proof: The following lemma will be used in the proof of this theorem.

Lemma 2.3. Let P be a finitely additive set functions on $\mathcal{F}$ and $P : \mathcal{F} \to [0, 1]$ such that $P(\Omega) = 1$. For each $X \in \mathcal{X}$ there is a finitely additive set function Q_X on $\mathcal{G}$ such that the equality (1.7) is satisfied and such that the restriction of Q_X to $\mathcal{F}$ is equal to P.

Proof: Define $\tilde{\rho}(X) = E_P[\rho_{\mathcal{F}}(X)]$, then $\tilde{\rho}(X)$ is a convex measure of risk. So for all $X \in \mathcal{X}$, there is a Q_X in $\mathcal{M}_{1,f}$ such that

$$\tilde{\rho}(X) = E_{Q_X}[-X] - \sup_{Y \in \mathcal{A}_{\tilde{\rho}}} E_{Q_X}[-Y] \qquad (1.10)$$

and for all $Z \in \mathcal{X}$

$$\tilde{\rho}(Z) \geq E_{Q_X}[-Z] - \sup_{Y \in \mathcal{A}_{\tilde{\rho}}} E_{Q_X}[-Y] \qquad (1.11).$$

Since X is bounded, from the equality (1.10), $\sup_{Y \in \mathcal{A}_{\tilde{\rho}}} E_{Q_X}[-Y]$ is a real number. It will be denoted by $\alpha(Q_X)$. If $Z = \beta I_B$ for all $\beta \in \mathbb{R}$ in (1.11),

$$E_P[\rho_{\mathcal{F}}(\beta I_B)] \geq E_{Q_X}[-\beta I_B] - \alpha(Q_X).$$

From multiplicative invariance $\rho_{\mathcal{F}}(\beta I_B) = \rho_{\mathcal{F}}(\beta) I_B$. Then, $\rho_{\mathcal{F}}(\rho_{\mathcal{F}}(\beta) + \beta) = 0$. Since $\rho(0) = 0$, by monotonicity $\rho(0) = -\beta$. Therefore

$$E_P[-\beta I_B] \geq -\beta Q_X(B) - \alpha(Q_X),$$

$$P(B) - \beta \geq -\beta Q_X(B) - \alpha(Q_X).$$

As a result, $0 \geq \alpha(Q_X) \geq \beta(P(B) - Q_X(B))$ for all $\beta \in \mathbb{R}$ and $B \in \mathcal{F}$. This inequality is satisfied for all β only if $P(B) = Q_X(B)$ for all B. This means that the restriction of Q_X to $\mathcal{F}$ is equal to P. Then (1.10) can be written as

$$E_{Q_X}[\rho_{\mathcal{F}}(X)] = E_{Q_X}[-X] - \sup_{Y \in \mathcal{A}_{\tilde{\rho}}} E_{Q_X}[-Y].$$

Since $\mathcal{A}_{\rho_{\mathcal{F}}}$ is contained in $\mathcal{A}_{\tilde{\rho}}$,

$$E_{Q_X}[\rho_{\mathcal{F}}(X)] \leq E_{Q_X}[-X] - \sup_{Y \in \mathcal{A}_{\rho_{\mathcal{F}}}} E_{Q_X}[-Y].$$

Moreover, from theorem (1.5), the converse inequality also holds. Then

$$E_{Q_X}[\rho_{\mathcal{F}}(X)] = E_{Q_X}[-X] - \sup_{Y \in \mathcal{A}_{\rho_{\mathcal{F}}}} E_{Q_X}[-Y]. \qquad (1.12)$$

Now assume that there is an $\mathcal{F}$ measurable set B such that inequality (1.8) is strict for Q_X. There is a $Y_0 \in \mathcal{A}_{\rho_{\mathcal{F}}}$ such that

$$E_{Q_X}[\rho_{\mathcal{F}}(X) I_B] > E_{Q_X}[-X I_B] - E_{Q_X}[-Y_0 I_B].$$

Let $Y = Y_0 I_B + (X + \rho_{\mathcal{F}}(X)) I_{\Omega - B}$. From bifurcation property $Y \in \mathcal{A}_{\rho_{\mathcal{F}}}$. Furthermore

$$\begin{aligned} E_{Q_X}[\rho_{\mathcal{F}}(X) I_B] &> E_{Q_X}[-X I_B] - E_{Q_X}[-Y I_B] - E_{Q_X}[-Y I_{\Omega-B}] + E_{Q_X}[-X I_{\Omega-B}] \\ &+ E_{Q_X}[\rho_{\mathcal{F}}(X) I_{\Omega-B}] \end{aligned}$$

This contradicts (1.12). So Q_X satisfies equality (1.7) for all $B \in \mathcal{F}$ and the restriction of Q_X to $\mathcal{F}$ is equal to P.

Con't of the proof: Let P be a probability measure on $(\Omega, \mathcal{F})$. Let $X \in \mathcal{X}$; from lemma 1.3 there exists a finitely additive set function Q_X such that equality (1.7) is satisfied for all $B \in \mathcal{F}$ and the restriction of Q_X to $\mathcal{F}$ is equal to P. It remains to prove that Q_X is a probability measure on $(\Omega, \mathcal{G})$.

Let $(A_n)_{n \in \mathrm{N}}$ be an increasing sequence in $\mathcal{G}$ and $\bigcup_{n \in \mathrm{N}} A_n = \Omega$. Then it must be proved that $Q_X(A_n)$ converges to 1. Applying equality (1.7) to $B = \Omega$;

$$E_P[\rho_{\mathcal{F}}(X)] = E_{Q_X}[-X] - \alpha(Q_X)$$

where $\alpha(Q_X) = \sup_{Y\in\mathcal{A}_{\rho_{\mathcal{F}}}} E_{Q_X}[-Y]$ and X and $\rho_{\mathcal{F}}(X)$ are bounded. Therefore $\alpha(Q_X)$ is finite.

Let $\lambda > 0$. Apply inequality (1.8) to λI_{A_n},

$$E_{Q_X}[\lambda I_{A_n}] \geq -E_P[\rho_{\mathcal{F}}(\lambda I_{A_n})] - \alpha(Q_X).$$

As n tends to infinity, $\rho_{\mathcal{F}}(\lambda I_{A_n})$ tends to $\rho_{\mathcal{F}}(\lambda) = -\lambda$. Therefore

$$\begin{aligned} \liminf_{n\to\infty} \lambda E_{Q_X}[I_{A_n}] &\geq \lim_{n\to\infty}(-E_P[\rho_{\mathcal{F}}(\lambda I_{A_n})] - \alpha(Q_X)), \\ \liminf_{n\to\infty} E_{Q_X}[I_{A_n}] &\geq 1 - \frac{\alpha(Q_X)}{\lambda}. \end{aligned}$$

As λ goes to infinity, $\liminf_{n\to\infty} E_{Q_X}[I_{A_n}] = \liminf_{n\to\infty} Q_X(A_n) \geq 1$. This ends the proof.

Convex risk measures conditional to $\mathcal{F}$ have representations in the following form, when they are continuous from below.

$$\rho_{\mathcal{F}}(X) = \inf_{g\in\varepsilon_{\mathcal{F}}}\{\forall Q \in \mathcal{M}_1(\Omega,\mathcal{G}), g \geq E_Q[-X \mid \mathcal{F}] - esssup_{Y\in\mathcal{A}_{\rho_{\mathcal{F}}}} E_Q[-Y \mid \mathcal{F}] \;\; Q\; a.s.\}$$

,where $M_1(\Omega,\mathcal{G})$ is the set of probability measures on $(\Omega,\mathcal{G})$. Under such a representation the penalty function equals to

$$\alpha(Q) = esssup_{Y\in\mathcal{A}_{\rho_{\mathcal{F}}}} E_Q[-Y \mid \mathcal{F}] \;\; Q\; a.s.$$

2.3.2 Convex Conditional Risk Measures Under Partial Uncertainty

In this part again Ω represents the infinite set of possible scenarios. A financial position is a bounded map on this set. $\mathcal{X}$ is the linear space of financial positions. σ algebra $\mathcal{F}$ represents all the accessible information for the investor. $\varepsilon_{\mathcal{F}}$ is the set of all bounded real valued $(\Omega,\mathcal{F})$ measurable maps. Unlike the previous part, there exists a probability measure P on $\mathcal{F}$. The aim is to define a risk measure conditional to $(\Omega, \mathcal{F}, P)$.

Definition 2.13. A mapping $\rho_{\mathcal{F}} : \mathcal{X} \to \mathrm{L}^{\infty}(\Omega, \mathcal{F}, P)$ is called a risk measure conditional to probability space $(\Omega, \mathcal{F}, P)$, if it satisfies the following conditions.

1. Monotonicity: For all $X, Y \in \mathcal{X}$; if $X \leq Y$, then $\rho_{\mathcal{F}}(X) \geq \rho_{\mathcal{F}}(Y)$ P $a.s.$

2. Translation Invariance: For all $Y \in \varepsilon_{\mathcal{F}}$ and for all $X \in \mathcal{X}$; $\rho_{\mathcal{F}}(X+Y) = \rho_{\mathcal{F}}(X) - Y$ P $a.s.$

3. Multiplicative Invariance: For all $X \in \mathcal{X}$ and for all $A \in \mathcal{F}$; $\rho_{\mathcal{F}}(XI_A) = I_A \rho_{\mathcal{F}}(X)$ P $a.s.$

Definition 2.14. A risk measure defined on $\mathcal{X}$, conditional to the probability space $(\Omega, \mathcal{F}, P)$ is called convex if for all $X, Y \in \mathcal{X}$ and for all $\lambda \in [0, 1]$;

$$\rho_{\mathcal{F}}(\lambda X + (1-\lambda)Y) \leq \lambda \rho_{\mathcal{F}}(X) + (1-\lambda)\rho_{\mathcal{F}}(Y) \ Pa.s.$$

Definition 2.15. The $\mathcal{F}$ acceptance set of a risk measure conditional to probability space $(\Omega, \mathcal{F})$ is

$$\mathcal{A}_{\rho_{\mathcal{F}}} = \{X \in \mathcal{X} \mid \rho_{\mathcal{F}}(X) \leq 0 \ Pa.s.\}.$$

Proposition 2.8. Let $\rho_{\mathcal{F}}$ be a risk measure conditional to the probability space $(\Omega, \mathcal{F}, P)$ with acceptance set $\mathcal{A} = \mathcal{A}_{\rho_{\mathcal{F}}}$. Then $\mathcal{A}$ satisfies the properties 1 and 2 of proposition 1.7 ($\mathcal{A}$ is a closed non-empty set satisfying hederitary and bifurcation properties). Furthermore it satisfies
3. Positivity: Every $\mathcal{F}$ measurable element of $\mathcal{A}$ is positive $Pa.s.$
4. $\rho_{\mathcal{F}}$ can be recovered from $\mathcal{A}$;

$$\rho_{\mathcal{F}}(X) = essinf\{Y \in \varepsilon_{\mathcal{F}} \mid X + Y \in \mathcal{A}\}.$$

Proof: 3. As an easy consequence of translation invariance and multiplicative invariance the restriction of any risk measure conditional to $(\Omega, \mathcal{F}, P)$, to $\varepsilon_{\mathcal{F}}$ is equal to $-identity$ $Pa.s.$ As a result of this property for all $\mathcal{F}$ measurable X, $\rho_{\mathcal{F}}(X) = -X$ $Pa.s.$ and by the definition of $\mathcal{A}$ positivity follows.

4. Let $X \in \mathcal{X}$. Denote $B_X = \{f \in \varepsilon_{\mathcal{F}} \mid X + f \in \mathcal{A}\}$. Therefore $\rho_{\mathcal{F}}(X+f) \leq 0$ $Pa.s.$, by translation invariance $\rho_{\mathcal{F}}(X) = essinf\{Y \in \varepsilon_{\mathcal{F}} \mid X + Y \in \mathcal{A}\}$.

Definition 2.16. A convex risk measure is continuous from below if for all increasing sequence, X_n of elements of $\mathcal{X}$ converging to X, the decreasing sequence of $\rho_{\mathcal{F}}(X_n)$ converges to $\rho_{\mathcal{F}}(X)$ $Pa.s.$

Theorem 2.7. Let $\rho_{\mathcal{F}}$ be a convex risk measure conditional to probability space $(\Omega, \mathcal{F}, P)$. Assume that $\rho_{\mathcal{F}}$ is continuous from below, then for all $X \in \mathcal{X}$

$$\rho_{\mathcal{F}}(X) = essmax_{Q \in \mathcal{M}}(E_Q[-X \mid \mathcal{F}] - \alpha(Q))$$

,where $\alpha(Q) = esssup_{Y \in \mathcal{A}_{\rho_{\mathcal{F}}}} E_Q[-Y \mid \mathcal{F}]$. $\mathcal{M}$ is a set of probability measures on $(\Omega, \mathcal{G})$ where restriction to $\mathcal{F}$ is equal to P.

Proof: The line of reasoning is similar to the proof of theorem 1.1 and can be found in [BN04] pg 21.

Chapter 3

EXAMPLES OF COHERENT AND CONVEX RISK MEASURES

At the beginning, it was said that the reason for VaR to be so widely used is its simplicity in interpretation and easiness in application. If the main goal is finding a risk measure better than VaR, constructing perfect mathematical models, evaluating every possibility, would not mean much, if it cannot be used by the agents in the market. In the previous chapter it is assumed that an agent works with all possible distribution functions. The strong theoretical foundations given in the previous chapter are very important to ensure that we are on the right track. However, less complex, more practical specifications are needed to approximate the risk of a position when observed market data is the only input. Such examples will be given in this chapter. Some of these examples are still suffering from complications in application but others are strong rivals for VaR.

3.1 Expected Shortfall:

Like stated before, VaR only answers the question of what the maximum loss with $\alpha\%$ confidence level is. In other words, in α of hundred observations, a loss higher than VaR would not be observed. If this interpretation is rephrased, VaR is the minimum loss that an investor can face in $100 - \alpha$ days out of 100 [AT02a]. Value at Risk cannot tell the investor what he should expect under the worst scenarios.

Given that portfolio returns are represented by random variable X on probability space $(\Omega, \mathcal{G}, P)$, $E[X]$ denotes the expectation of X under the given probability distribution P; under the assumption that $E[X^-] < \infty$. P is the probability distribution of the historical returns. One of the answers to the above question is the Tail Conditional Expectation(TCE). Through the rest of this part VaR is defined as:

$$VaR_\alpha(X) = -\sup\{x \mid P(X \leq x) \geq \alpha\}.$$

Definition 3.1. Let X be a random variable representing the profit of the portfolio (positive values for profit and negative values for losses). For a determined confidence level α; TCE is defined as

$$TCE_\alpha(X) = -E\{X \mid X \leq -VaR_\alpha\}.$$

This measure gives us the mean loss under scenarios leading losses higher than or equal to VaR_α. But this measure is not coherent in general, since it fails to satisfy the subadditivity axiom as given in [D00]. Moreover, if the distribution is discontinuous, then $\{X \leq -VaR_\alpha\}$ can have probability higher than $(1-\alpha)$. Therefore, the outcome can be higher than our worst case expectation and does not answer our question. Such a situation is shown in Figure 3.1.

Expected shortfall (ES) is constructed as a solution for all problems stated above. [AT02a] explains it in the context of the following example. Take a large number of observations $\{x_i\}_{\{i=1,\dots,n\}}$ of random variable X. Sort them in ascending order as $x_{1:n} \leq x_{2:n} \leq \dots \leq x_{n:n}$. Then approximate number of $(1-\alpha)\%$ elements in the sample by $w = [n(1-\alpha)] = max\{m \mid m \leq n\alpha, m \in \mathrm{N}\}$. The set of $(1-\alpha)\%$ worst cases is represented by the least w outcomes $\{x_{1:n} \leq x_{2:n} \leq \dots \leq x_{w:n}\}$. Then a natural estimator for $1-\alpha$ quantile x^{α} is

$$x^{\alpha}_{(n)}(X) = x_{w:n}$$

and

$$ES^{\alpha}_{n}(X) = -\frac{\sum_{i=1}^{n} x_{i:n}}{w}$$

In order to understand the difference in this setting TCE^{α}_{n} is given as

$$TCE^{\alpha}_{n}(X) = -\frac{\sum_{i=1}^{n} x_i I_{\{x_i \leq x_{w:n}\}}}{\sum_{i=1}^{n} I_{\{x_i \leq x_{w:n}\}}}$$

Subadditivity of ES can be seen as

$$\begin{aligned} ES^{\alpha}_{n}(X+Y) &= -\frac{\sum_{i=1}^{n}(x+y)_{i:n}}{w} \\ &\leq -\frac{\sum_{i=1}^{n}(x_{i:n}+y_{i:n})}{w} \\ &= ES^{\alpha}_{n}(X) + ES^{\alpha}_{n}(Y). \end{aligned}$$

As can easily be seen in the Figure 3.2, giving the cumulative probability distribution of losses (i.e. positive values for the losses and negative values for the profits) when VaR is taken at 95 percent confidence level, for a discontinuous distribution ES uses only points $\{c, d\}$, while TCE is found through $\{a, b, c, d\}$. So TCE is always smaller than or equal to ES.

When the number of observations go to infinity, the result given above is extended by using the following definition.

Definition 3.2. Let X be a random variable and α be a specified probability

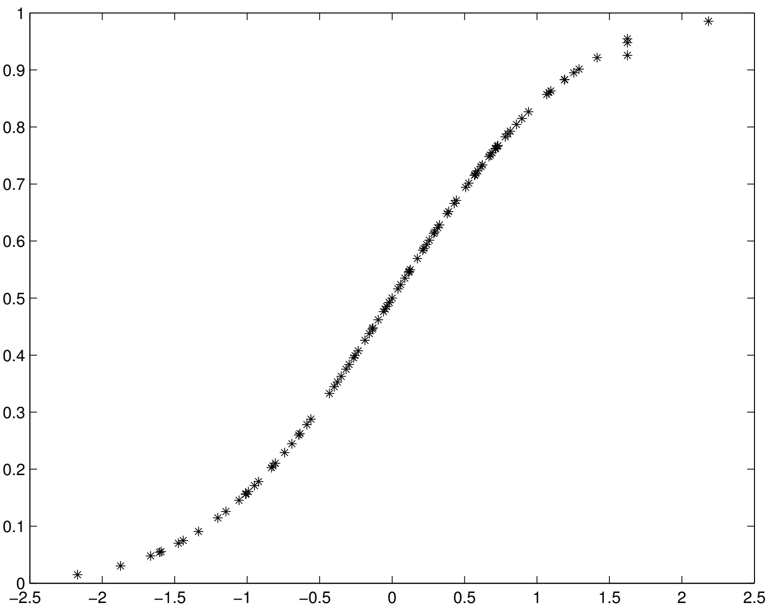

Figure 3.1: Loss distribution of a portfolio

level. Then ES is defined as;

$$ES_n^{\alpha}(X) = -\frac{1}{(1-\alpha)}(E[XI_{\{X \leq x^{\alpha}\}}] - x^{\alpha}(P(X \leq x^{\alpha})) - (1-\alpha))$$

,where $x^{\alpha} = lim_{n\to\infty} X_{w:n}$.

The first term on the right hand side is equal to TCE, the second term creates the difference between TCE and ES when the distribution is not continuous at $(1-\alpha)$ level. That is the exceeding part which has to be subtracted from TCE when $\{X \leq x^{\alpha}\}$ has a probability larger than $(1-\alpha)$, otherwise it disappears.

Remark: As it is discussed an insurance against the uncertainty of net worth of the portfolio, firm could or often would be regulated to hold an amount of riskless investment. This amount is named as risk capital. From a financial perspective this low return investment is a burden. So it is important to optimize this amount and fairly allocate it. To do this management should answer, how much of this risk capital is due to each of the departments. Answering this question risk capital can be allocated coherently. Moreover such an allocation

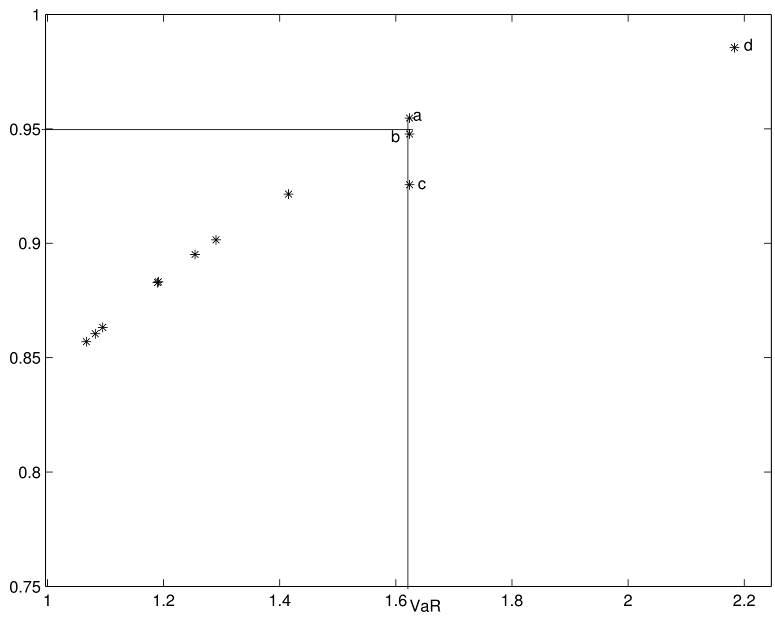

Figure 3.2: VaR and other related points in the loss distribution

provides a basis for the performance comparison.

In [We02] an allocation method for risk capital is said to be coherent if:

1. the risk capital is fully allocated to the portfolios, in particular,each portfolio can be assigned a percentage of the total risk capital,

2. no portfolio's allocation is higher than if they stood alone, (Similarly for any coalition of portfolios and coalition of fractional portfolios.)

3. a portfolio's allocation depends only its contribution to risk within the firm,

4. a portfolio that increases its cash position will see its allocated capital decreases by the same amount.

3.2 Worst Conditional Expectation:

TCE was not the only alternative offered by ADEH for $(1-\alpha)\%$ worst cases. As a coherent alternative, worst conditional expectation was also given.

Definition 3.3. The worst Conditional Expectation(WCE) at specified confidence level α is

$$WCE_\alpha(X) = -\inf\{E[X \mid A] : A \in \mathcal{A}, P(A) > 1-\alpha\}.$$

Although definition of WCE can be interpreted similar to ES, it is a little more complicated. WCE depends not only on the distribution of X but also on the σ algebra $\mathcal{G}$. Moreover it assumes that the investor has coplete information on the probability space. We are still working with the given probability distribution but now we also need to know all $\mathcal{A} \in \mathcal{G}$. For an infinite Ω this can be a little impractical.

3.3 Conditional Value at Risk:

Conditional Value at Risk (CVaR) is a coherent system constructed by Rockafeller and Uryasev, to quantify dangers beyond VaR. This structure is mainly an optimization problem, solved through linear programming techniques. For continuous distributions, CVaR at a given confidence level is the expected loss; given that loss is greater than or greater or equal to VaR at that level. Moreover if the distribution is continuous, TCE=ES=CVaR. $CVaR^+$ represents expected loss greater than VaR at level α and $CVaR^-$ represents expected loss greater than or equal to VaR at level α, equal to TCE. CVaR will be defined as a weighted average of $CVaR^+$ and VaR and $CVaR^- \leq CVaR \leq CVaR^+$.

Definition 3.4. When X is a random variable representing loss of a portfolio on the given probability space $(\Omega, \mathcal{G}, P)$,

$$CVaR = \frac{P(X \leq VaR^{\alpha}) - \alpha}{1-\alpha} VaR^{\alpha} + \frac{1 - P(X \leq VaR^{\alpha})}{1-\alpha} CVaR^{+}$$

or equivalently

$$\begin{aligned} CVaR &= \frac{1}{1-\alpha} \int_{\alpha}^{1} F^{-1}(u) du \\ &= \frac{1}{1-\alpha} \int_{F^{-1}(\alpha)}^{\infty} u dF(u) \end{aligned}$$

,where F^{-1} is the inverse distribution function of X.

Confirmation of the coherence of CVaR can be found in [P00].

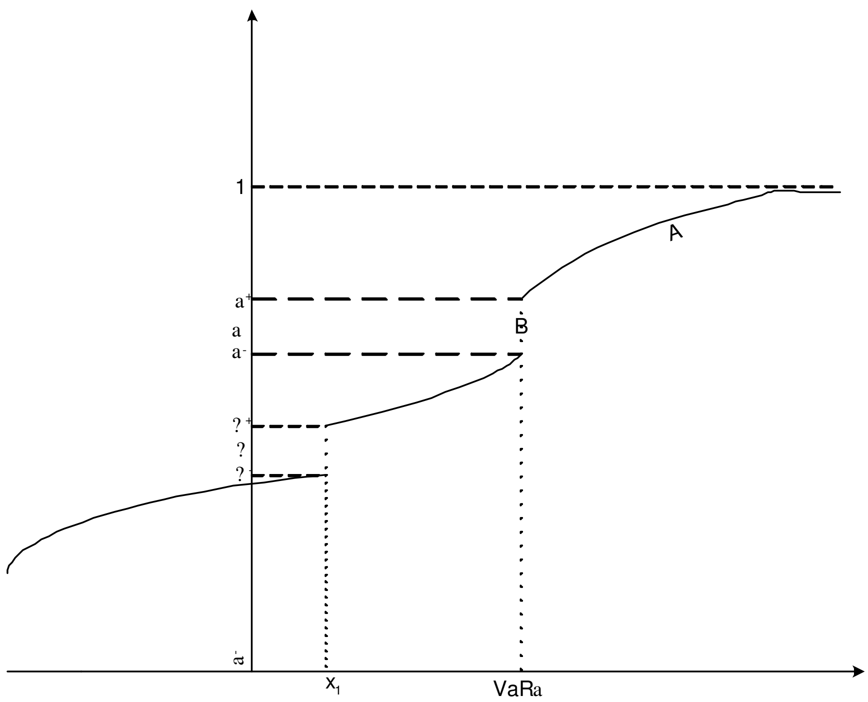

Figure 3.3: Loss distribution of portfolio X

In the figure below, showing the cumulative distribution function of the losses of portfolio X, there are two jumps. The first is at loss level x_1. A jump means that the probability of observing a loss of x_1 is equal to $(B^{+} - B^{-})$.

Value-at-Risk for the confidence level a is equal to VaR_a in the figure by definition. $CVaR^+$ is equal to the conditional expectation of the "A" part of the distribution. $CVaR^-$ is equal to the conditional expectation of part "A+B". $P(X \leq VaR^{\alpha}) - \alpha$ is $a^+ - a$. $1 - P(X \leq VaR^{\alpha})$ is $1 - a^+$. This means expectation of the part "A" is taken conditional to $(1-a)$ instead of $(1-a^+)$. The rest is taken from VaR_a.

Being a coherent risk measure CVaR has a convex acceptance set, which means every optimum is global. Therefore, it is generally used in portfolio optimization. When asset returns are given as $\varepsilon = (\varepsilon_1, ..., \varepsilon_k)$ for k assets and $x = (x_1, ..., x_k)$ is the amount of investment in those k assets, total portfolio return will be $-X = \varepsilon^T x$. The aim is minimizing CVaR under the given constraint (expected return of the portfolio exceeds a determined level). The problem can be formalized as follows;

$$\begin{aligned} &min_x && CVaR_{\alpha}(X) \\ &subject\ to && \varepsilon^T(X) \geq \mu \\ & && x^T 1 = 1 \\ & && x \geq 1 \end{aligned}$$

3.4 Coherent Risk Measures Using Distorted Probability

In all the examples given above, risk measures are constructed on the worst cases. But a good measure should consider the entire distribution. According to [W02]; considering only the worst cases creates a disincentive for risk management. When an agent uses ES, for instance, he not only ignores the distribution of losses less than a determined quantile but also cannot see extreme low frequency and high severity losses since ES is only a mean value. To overcome these deficiencies a risk measure, based on the mean value under distorted probability, will be offered.

In this setting, the end of the period loss is represented by X and it has cumulative distribution $F(x)$. In order to reflect real market conditions this distribution is assumed to be discontinuous.

Definition 3.5. Let $g : [0,1] \rightarrow [0,1]$ be an increasing function with $g(0) = 0$ and $g(1) = 1$. The transform $F^*(x) = g(F(x))$ defines a distorted probability distribution where g is called a distortion function.

Definition 3.6. A family of distortion risk measures using the mean-value under the distorted probability $F^*(x) = g(F(x))$

$$\rho(X) = E^*[X] = -\int_{-\infty}^{0} g(F(x))dx + \int_{0}^{\infty} (1 - g(F(x)))dx$$

is a coherent risk measure when g is a continuous mapping.

The first argument is the expected profit under F^* and the second argument is the mean of losses under F^*. Therefore all information contained in the original distribution can be used depending on "g".

After constructing the model, the crucial point is choosing the distribution function. [W02] recommends the following transformation.

$$g(u) = \phi(\phi^{-1}(u) - \lambda)$$

,where ϕ is a standard normal cumulative distribution. Then

$$F^*(x) = \phi(\phi^{-1}(F(x)) - \lambda)$$

Definition 3.7. For a loss variable X with distribution F, a new risk measure for capital requirement is defined as follows;

1. For a preselected confidence level α, let $\lambda = \phi^{-1}(x)$,
2. $F^*(x) = \phi(\phi^{-1}(F(x)) - \lambda)$,

3. Set the capital requirement to be the expected value under F^*

$$WT(\alpha) = E^*[X].$$

Example: Consider a portfolio X with the loss distribution given in the first two columns of the table below. For the confidence level 0.95, expected shortfall of this portfolio is 6.2. On the other hand, the risk measure based on the distorted probabilities given in the third column, again at 0.95 confidence level, is 5.36. (For this example WT is smaller than ES, but this can change for different distributions.) Furthermore, for comparison, in Figure 3.3 both $f(x)$ and $f^*(x)$ are shown. In the figure f^* is represented by * and for f $\diamondsuit$ is used.

x	$f(x)$	$f^*(x)$
-5.5	0.01	0.00010507
-5	0.05	0.0016707
-4	0.09	0.0076707
-3	0.13	0.023353
0	0.19	0.063602
1	0.21	0.13966
2	0.15	0.21998
4	0.11	0.2781
5	0.04	0.25684
7	0.02	0.0090228

3.5 Convex Risk Measures Based on Utility

Until now coherent risk measures depending only on probability distributions have been discussed. In this section risk will be measured also in terms of subjective utility functions. Convex risk measures based on utility are formed by Föllmer *et. al.* They say that when X represents the returns of a given portfolio,

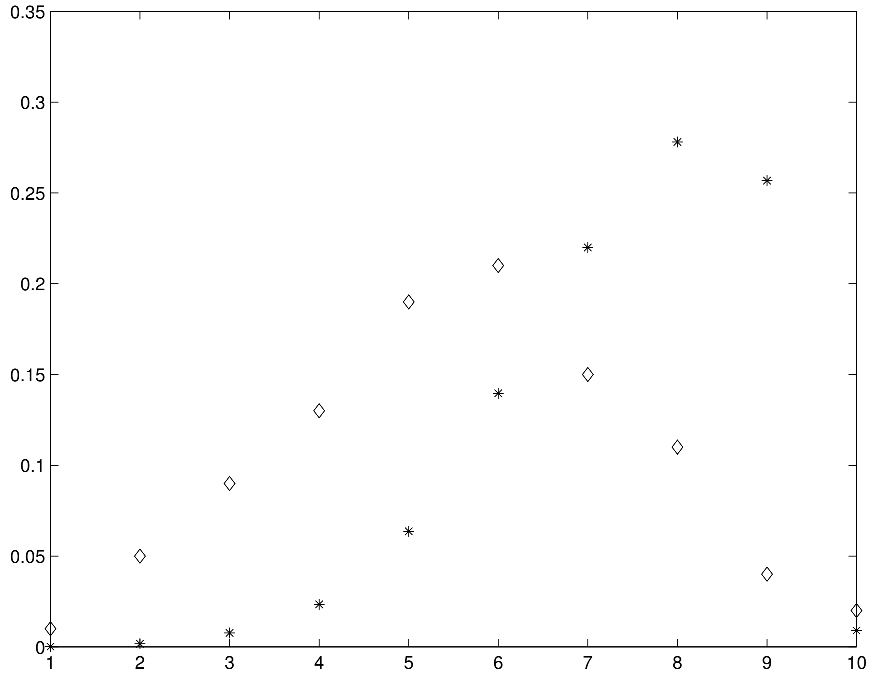

Figure 3.4: Comparison of $f(x)$ with $f^*(x)$

from an investors point of view, risk is related to losing money. Therefore it is related to X^-. The risk of facing a loss is called as shortfall risk. In order to find this risk, a loss function is used. This is a function giving the amount of disutility that is caused by a given loss level.

Definition 3.8. A function $l : \mathbb{R} \rightarrow \mathbb{R}$ is called a loss function if it is increasing, not identically constant and has the form

$$l(x)) = -u(-x)$$

,where $u(x)$ is an increasing, concave and continuous utility function.

It is assumed that the investor knows his loss function and in the given probability space $(\Omega, \mathcal{G}, P)$ he measures the risk of a position with the expected loss. Then it is natural to define the set of acceptable positions with a predetermined disutility level, x_0.

$$\mathcal{A} = \{X \in \mathcal{X} \mid E[l(-X)] \leq x_0\}$$

Using this acceptance set a convex risk measure, having the representation (1.3), can be formalized. An example, given in [FS02b], is discussed below.

Example: If the loss function is given as $l(X) = e^x$ and $x_0 = 1$,

$$\rho(X) = inf\{m \in \mathrm{R} \mid E_P[e^{-m-X} \leq 1]\}.$$

To find the infimum take

$$\begin{aligned} E_P[e^{-m-X}] &= 1, \\ e^{-m}E_P[e^{-X}] &= 1, \\ log\ E_P[e^{-X}] &= \log(e^m). \end{aligned}$$

Therefore $\rho(X) = \log\ E_P[e^{-X}]$. For a convex risk measure given as

$$\rho(X) = \sup_{Q \in \mathcal{M}_1(P)} (E_Q[-X] - \alpha(Q)),$$

$\alpha(Q) \geq \sup_{X \in \mathcal{X}}(E_Q[-X] - \rho(X)) \geq \alpha_{min}(Q)$ from *Theorem 1.4*, where $\mathcal{M}_1(P)$ is the set of probability measures that are absolutely continuous with respect to P. When $\alpha(Q) = \alpha_{min}(Q)$, the above inequality becomes an equality. As a result

$$\alpha_{min}(Q) = sup_{X \in \mathcal{X}}(E_Q[-X] - \log\ E_P[e^{-X}]).$$

In order to identify $\alpha_{min}(Q)$, the supremum of the following must be found.

$$\int -X\frac{dQ}{dP}dP - \log \int e^{-X}dP$$

Suppose $Y \in \mathrm{L}^2$

$$\int -(X + \delta Y)\frac{dQ}{dP}dP - \log \int e^{-(X+\delta Y)}dP$$

Differentiate with respect to Y

$$\int -Y\frac{dQ}{dP}dP - \frac{1}{\int e^{-(X+\delta Y)}dP}\int -Ye^{-(X+\delta Y)}dP = 0$$

$$\int -Y[\frac{dQ}{dP} - \frac{1}{\int e^{-X}dP}e^{-X}]dP = 0$$

The above equation is satisfied for all Y's if the argument inside the brackets is 0.

$$\frac{dQ}{dP} - \frac{1}{\int e^{-X}dP}e^{-X} = 0$$

$$\frac{dQ}{dP} = \frac{e^{-X}}{\int e^{-X}dP}$$

And it follows that

$$\log\int e^{-X}dP = -X - \log\frac{dQ}{dP}$$

When expectation with respect to Q is taken,

$$E_P[\frac{dQ}{dP}\log\frac{dQ}{dP}] = E_Q[-X] - \log\ E_P[e^{-X}] = \alpha_{min}(Q)$$

This means our minimal penalty function is defined in terms of relative entropy (Appendix, A8). For each loss function, following the above way, minimal penalty function can be found. Also there exists a common formulation.

Theorem 3.1. For any convex loss function l, the minimal penalty function is given by

$$\alpha_{min}(Q) = \inf_{\lambda>0}\frac{1}{\lambda}\big(x_0 + E[l^*(\lambda\frac{dQ}{dP})]\big)$$

,where $Q \in \mathcal{M}_1(P)$ and l^* is the conjugate function i.e. $l^*(z) = \sup_x \in \mathrm{R}(zx - l(X))$.

The proof of this theorem can be found in [FS02b].

Example: Take

$$l(x) = \begin{cases} \frac{1}{p}x^p \text{ if x} \geq 0 \\ 0 \text{ otherwise} \end{cases}$$

where $p > 1$. Then

$$l^*(z) = \begin{cases} \frac{1}{q} z^q \text{ if x} \geq 0 \\ +\infty \text{ otherwise} \end{cases}$$

For a given $x_0 > 0$

$$\begin{aligned} \alpha_{min}(Q) &= \inf_{\lambda > 0} \frac{1}{\lambda} \big(x_0 + E[\frac{1}{q} (\lambda \frac{dQ}{dP})^q] \big) \\ &= \frac{x_0}{\lambda} + \int \frac{1}{q} \lambda^{q-1} (\frac{dQ}{dP})^q dP \qquad (2.1) \end{aligned}$$

Let $\varphi = \frac{dQ}{dP}$ and differentiate (2.1).

$$-\frac{x_0}{\lambda^2} + (q-1)\lambda^{q-2} \int \frac{1}{q} \varphi^q dP = 0$$

As a result , the infimum is attained for

$$\lambda_Q = \big(x_0 \frac{q}{q-1} E_P[(\frac{dQ}{dP})^q] \big)^{\frac{1}{q}}.$$

Chapter 4

MULTI PERIOD COHERENT RISK MEASURES

In the first two chapters the financial decision process is based on one period models. However, generally agents do not see an investment decision as a one period action. Considering this, another approach, treating risk measurement as a dynamic process consisting of risk evolving over several periods of uncertainty, was developed. In this kind of setting the availability of new information through the process makes revaluation in locked-in positions possible. In other types of positions intermediate actions such as cash inflows or outflows is another option. For instance when intermediate action is a possibility, insolvency time through the process can be determined and by the injection of necessary amount of cash this liquidity problem can be solved. This kind of liquidity injection strategy works, if there are more favorable dates and events after this insolvency time. On the other hand, if agents are working in locked-in positions, i.e. investment decisions are given at the beginning, as time passes, positions are revaluated in the light of the new information. As a result in the decision phase this revaluation process should also be considered by the agents. As such, it can be understood that multiperiod risk measurement is slightly different from one period models. In this chapter the multiperiod decision process is investigated in terms of coherent risk measures. For this, we will first present the necessary assumptions and notations. Following this, dynamic consistency, a necessary feature of dynamic risk measurement, will be investigated. The chapter will end with an example.

4.1 Assumptions and Notations

Working with an elementary model it is assumed that:

1) Agents are working with locked in positions.

2) There are T periods of uncertainty; working in discrete, time date is shown as $t = 0, 1, \ldots, T$.

3) Ω represents the finite set of states of the world.

4) A denotes a finite set of events at each t. (Such a setting makes it easier to use trees, for instance, in a binomial tree model $A = \{up,\ down\}$)

5) Ω is the set of all sequences $(\alpha_1, \ldots \alpha_T)$, $\alpha_i \in A$. These sequences are called full histories.

6) A collection of sequences $(\alpha_1, \ldots \alpha_\tau)$ of length τ $(1 \leq \tau \leq T)$ compose Ω_τ. $\Omega'_\tau = \bigcup_{1 \leq \tau \leq T} \Omega_\tau$ is the set of all sequences of length at most τ.$(\Omega'_\tau = \Omega')$ $\Omega'' = \Omega'_{T-1}$ Both Ω'' and Ω' are called partial histories. Ω_0 represents the initial state of the economy. The situation at the initial time is represented by the sequence of zero length which consists of a single element 0.

7) For $\omega = (\alpha_1, \ldots \alpha_T) \in \Omega$; the τ restriction of ω is $\omega_\tau = (\alpha_1, \ldots \alpha_\tau)$. If $\omega' \in \Omega'$ is ancestor of $\omega \in \Omega$ i.e. $\omega' = \omega_\tau$, then it is said that $\omega' \preceq \omega$.

8) All possible full histories following the partial history $\omega' \in \Omega'$ are represented by

$$F(\omega') = \{\omega \in \Omega \mid \omega' \preceq \omega.\}$$

From another point of view F_t represents observed history up to the time t.

9) $\mathcal{F}_\tau$ is the σ algebra generated by the sets $F(\omega_\tau)$ with $\omega_\tau \in \Omega_\tau$.

10) $\mathcal{X}$ denotes the set of possible portfolios. For each $X \in \mathcal{X}$, $X : \Omega \to \mathbb{R}$. However, this time X must be seen as a value process, $X = (X_t)_{0 \leq t \leq T}$. Moreover X_t is an F_t-adapted process.

Remark: $X = (X_t)_{0 \leq t \leq T}$ is a function on the product space $\{0, 1, \ldots, T\} \times \Omega$ and it is constant for each date in any full history having the same partial history ω_t.

$$X_t(\omega^1) = X_t(\omega^2) \text{ as long as } \omega^1 \text{ and } \omega^2 \in F_t.$$

In this chapter instead of risk measure ρ, risk adjusted value ϕ will be used. It's definition will again be given in terms of acceptable positions.

Definition 4.1. A coherent acceptance set of value processes is a closed convex cone of $\mathcal{X}$ with vertex at the origin, containing positive orthant and intersecting the negative orthant only at the origin.

Definition 4.2. A coherent risk adjusted value associated to a coherent cone $\mathcal{A}_{cc}$ is

$$\phi(X) = \sup\{m \mid X - m \in \mathcal{A}_{cc}\}.$$

As can easily be seen, risk adjusted value is only the negative of risk measure ρ. It gives the largest amount of capital that can be subtracted from the position and still leave it acceptable. Such a conversion is used to simplify calculations.

In [Rie04], a dynamic risk adjusted value $\phi = (\phi_t)_{t=0,\ldots,T}$, where $\phi_t : (X) \times \Omega \to \mathbb{R}$, is said to satisfy

1. Independence of the past: For all $X, Y \in \mathcal{X}$ and for all $t \in [0, T]$; if $X_s(\omega) = Y_s(\omega)$ for all $s \geq t$ and for all $\omega \in \Omega$, then $\phi_t(X) = \phi_t(Y)$.

2. Monotonicity: For $X, Y \in \mathcal{X}$; if $X \geq Y$ then $\phi(X) \geq \phi(Y)$.

3. Translation Invariance (with respect to predictable income streams): If z is an F_t measurable constant cash inflow at date $\tau \geq t$, then

$$\phi_t(X + z) = \phi_t(X) + \frac{z}{(1+r)^{\tau - t}}$$

,where r is the risk free rate of return. (In our setting it is assumed to be 0).

If it is assumed that bygones are bygones, it is natural that past payments should not affect the risk level of future payments. In multiperiod setting risk adjusted value is translation invariant with respect to the predictable future cash additions. Adding amount z at date τ for all ω_τ is equal to adding present the value of z to the position today.

Moreover, a dynamic risk adjusted value is coherent if it satisfies

1. Homogeneity: For all $X \in \mathcal{X}$ and $t = 0, \ldots, T$; $\phi_t(\lambda X) = \lambda \phi_t(X)$, where $\lambda \geq 0$.
2. Superadditivity: For all $X, Y \in \mathcal{X}$ and $t = 0, \ldots, T$; $\phi_t(X+Y) \geq \phi_t(X) + \phi_t(Y)$.

4.2 Dynamic Consistency and Recursivity of Dynamic Coherent Risk Adjusted Values

In a multiperiod setting the acceptability of a position should not be considered simply as a function of the position itself but also as a function of the available information. Therefore, the decision process must be consistent in terms of information flow. That means that if a position is acceptable in period t+1, then it must also be acceptable on date t, since information set at t+1 already includes the information provided by F_t. This property is summarized under the heading 'dynamic consistency' as follows.

Definition 4.3. A coherent risk adjusted value defined on $\mathcal{X} \times \Omega$ is said to be dynamically consistent if

$$\phi_{t+1}(X) = \phi_{t+1}(Y) \; for \; all \; \omega_{t+1} \; then \; \phi_t(X) = \phi_t(Y)$$

for all $t = 0, \dots, T$ and all positions $X, Y \in \mathcal{X}$.

Dynamic consistency is provided by a consistent set of probability measures. Consistency in terms of probability measures are related being a product type probability measure which will be introduced below. However in order to define the consistency of probability measures, we will firstly give the necessary notations.

Let P be any probability measure on $(\Omega, \mathcal{F}_T)$ assigning $P(\omega) \geq 0$ for each $\omega \in \Omega$. Then the marginal probability of a sequence $\omega' \in \Omega'$ is

$$P(\omega') = \sum_{\omega' \preceq \omega} P(\omega) = P(F(\omega')). \qquad (A)$$

(In a tree framework this is the probability of getting to the node representing ω'.) The conditional probability of realization of $\omega \in \Omega$, given a sequence $\omega' \in \Omega'$ such that $\omega \succeq \omega'$, is

$$P(\omega \mid \omega') = \frac{P(\omega)}{P(\omega')}.$$

A single period density (i.e. probability of ω' followed by α in the next period) is

$$P^s(\alpha \mid \omega') = \frac{P(\omega'\alpha)}{P(\omega')}. \qquad (B)$$

Then

$$\begin{aligned} P(\omega \mid \omega') &= \frac{P(\omega'\alpha)}{P(\omega')} \frac{P(\omega)}{\omega'\alpha} \\ &= P^s(\alpha \mid \omega') P(\omega \mid \omega'\alpha). \end{aligned}$$

The conditional expectation is

$$\begin{aligned} E(X \mid \omega') &= \sum_{\omega \succeq \omega'} P(\omega \mid \omega')X(\omega) \\ &= \sum_{\alpha \in A} P^s(\alpha \mid \omega') \sum_{\omega \succeq \omega'} P(\omega \mid \omega'\alpha)X(\omega) \\ &= E_{P^s_{\omega'}} E[X \mid \omega'\alpha]. \end{aligned}$$

Using all of these, given a collection $\mathcal{P}$ of probability measures on Ω,

$$\phi_t(X) = \inf_{P \in \mathcal{P}} E_P[X \mid \omega_t]. \qquad (3.1)$$

When risk adjusted value is defined as equation (3.1), it is easy to verify that it satisfies the conditions of coherency.

If $\mathcal{P}$ is given as a collection of probability measures on Ω , for each partial history $\omega' \in \Omega''$ a collection of single period probability measures can be found by

$$\mathcal{P}^s(\omega') = \{P^s(\cdot \mid \omega') \mid P \in \mathcal{P} \ with \ P(\omega') > 0\}.$$

Conversely if for each $\omega' \in \Omega''$ a collection of single period measure $\mathcal{P}^s(\omega')$ is given, then the family $\{\mathcal{P}^s(\omega')\}_{\omega' \in \Omega''}$ defines a collection of probability measures on Ω by

$$\mathcal{P} = \{P \mid P^s(\cdot \mid \omega') \in \mathcal{P}^s(\omega') \ for \ all \ \omega' \in \Omega'' \ such \ that \ P(\omega') > 0\}.$$

This generated probability measure has the form

$$P : (\alpha_1, \ldots, \alpha_T) \to \prod_{t=1}^{T} P^s_t(\alpha_t).$$

Starting with a given set of probability measures, $\mathcal{P}$ is the same as the probability collection gathered by firstly finding single-period probability measures and then generating probability measures as described above. This equivalence is valid if $\mathcal{P}$ is a collection of product type probability measures.

Definition 4.4. A collection of probability measures $\mathcal{P}$ on Ω is said to be a product type if

$$\mathcal{P} = \{P \mid P^s(\cdot \mid \omega') \in \mathcal{P}^s_{\omega'} \; for \; all \; \omega' \in \Omega'' \; such \; that \; P(\omega') > 0.\}$$

(In [ADEHKu02] this product type construction of probability measures is given as the "Stability by Pasting" property.) This property means that any composition of single period probability measures is still in $\mathcal{P}$.

Product type measures are closely related to the recursivity of the risk measurement. By recursivity it is meant that

$$\phi(X \mid \omega_t) = \inf_{P \in \mathcal{P}} E_P[\phi(X \mid \omega_{t+1}) \mid \omega_t]$$

Actually working with product type probability measures is equivalent to the recursivity property of the risk measure.

$$\phi_t(X) = \inf_{P \in \mathcal{P}} E_P[X \mid \omega_t]$$

Using product type measures, this conditional expectation is minimized by finding the single period probability measures minimizing expected value at each date. Afterwards these single period measures are combined to reach minimizing probability measure. The solution offered by the recurrence relation is exactly the same: step by step resolution of the uncertainty.

For the coherent multiperiod acceptability measures the properties of dynamic consistency and representability by a product type probability measures are equivalent. Representability by product type probability measures for a risk adjusted value in equation (3.1) means that it has the property of dynamic consistency. Because if a risk measure can be represented by product type probability measures, it has the recursivity property. This means that acceptability of X_t, given the history up to t ω_t, is determined in terms of values of

$\phi(X \mid \omega_{t+1})$. This also the main argument of dynamic consistency.

4.3 An Example

In this section a simple theoretical example will be explored in order to illustrate the notations and definitions of the first two sections. Our three period ($t = 0, 1, 2, 3$) model will be given as a trinomial tree. For the investigated portfolio X, there are three possible movements in each period. This means

$$A = \{up, middle, down\}.$$

Then the set of all possible sequences are given as
$\Omega = \{uuu, uum, uud, umu, umm, umd, udu, udm, udd, muu, mum, mud, mmu, mmm, mmd, mdu, mdm, mdd, duu, dum, dud, dmu, dmm, dmd, ddu, ddm, ddd\}$
Our random variable X, which is the portfolio value process, is represented by a tree in the Figure 4.1. Since we are working with locked-in positions, it is assumed that the investor cannot take an intermediate action. Therefore our main concern is the final portfolio values at $t = 3$. Only a revaluation of the portfolio by using the risk adjusted value process is allowed. The final portfolio values are given in the table below.

$X_3(uuu)$	$X_3(uum)$	$X_3(uud)$	$X_3(umu)$	$X_3(umm)$
23634,75	23516,34	23472,44	23464,46	23437,79
$X_3(umd)$	$X_3(udu)$	$X_3(udm)$	$X_3(udd)$	$X_3(muu)$
23293,21	23215,57	23176,7	23150,06	23132,59

$X_3(mum)$ 23075,58	$X_3(mud)$ 23049,51	$X_3(mmu)$ 23008,07	$X_3(mmm)$ 23006,45	$X_3(mmd)$ 22943,67
$X_3(mdu)$ 22931,37	$X_3(mdm)$ 22799,16	$X_3(mdd)$ 22679,88	$X_3(duu)$ 22625,44	$X_3(dum)$ 22618,03
$X_3(dud)$ 22615,99	$X_3(dmu)$ 22566,4	$X_3(dmm)$ 22560,87	$X_3(dmd)$ 22544,28	$X_3(ddu)$ 22486,2
$X_3(ddm)$ 22168,87	$X_3(ddd)$ 22104,69			

Table 4.1: Final portfolio values

It is assumed that agents foresee only three possible scenarios or, in other words, three possible probability distributions. First, one assumes all states have an equal probability of 0.0370; second, one assumes asset values will be distributed normally and, according to the last scenario, asset prices will be determined by the t distribution. Using the formula (A), probability distributions on Ω' can also be found. Moreover, using formula (B), all of the single period probability measures can be found. For instance

$$P_2(u \mid u) = \frac{P_2(uu)}{P_2(u)} = \frac{0.0411}{0.2321} = 0.1771$$

It is easier to see these single period probability measures on a trinomial tree. Single period probabilities for the normal distribution are shown in Figure 4.2. When this process is repeated for the other two distributions, single period probability measures for P_1 and P_3 are calculated. Trees for these two distribu-

	P_1	P_2	P_3
$P(uuu)$	0.0370	0.0047	0.0057
$P(uum)$	0.0370	0.0135	0.0079
$P(uud)$	0.0370	0.0230	0.0218
$P(umu)$	0.0370	0.0248	0.0358
$P(umm)$	0.0370	0.0249	0.0378
$P(umd)$	0.0370	0.0286	0.0397
$P(udu)$	0.0370	0.0349	0.0453
$P(udm)$	0.0370	0.0388	0.0482
$P(udd)$	0.0370	0.0391	0.0493
$P(muu)$	0.0370	0.0478	0.0505
$P(mum)$	0.0370	0.0483	0.0555
$P(mud)$	0.0370	0.0484	0.0560
$P(mmu)$	0.0370	0.0497	0.0562
$P(mmm)$	0.0370	0.0500	0.0569
$P(mmd)$	0.0370	0.0501	0.0573
$P(mdu)$	0.0370	0.0503	0.0562
$P(mdm)$	0.0370	0.0504	0.0492
$P(mdd)$	0.0370	0.0502	0.0430
$P(duu)$	0.0370	0.0500	0.0409
$P(dum)$	0.0370	0.0496	0.0391
$P(dud)$	0.0370	0.0459	0.0379
$P(dmu)$	0.0370	0.0424	0.0344
$P(dmm)$	0.0370	0.0397	0.0331
$P(dmd)$	0.0370	0.0357	0.0232
$P(ddu)$	0.0370	0.0261	0.0161
$P(ddm)$	0.0370	0.0207	0.0024
$P(ddd)$	0.0370	0.0126	0.0006

Table 4.2: Possible probability distributions

tions are given in Appendix A9.

Under the assumption that we are working with product type probability measures, any combination of these single period probability measures are still in $\mathcal{P}$. Every point in the tree represents a node. Except the last nodes every node has a single period probability distribution, giving the conditional probability distribution for that node. So, for instance, the single period probability distribution for node u can be take from P_1, for another node d, a single period

	P_1	P_2	P_3
$P(uu)$	0.1110	0.0411	0.0354
$P(um)$	0.1110	0.0782	0.1132
$P(ud)$	0.1110	0.1127	0.1429
$P(mu)$	0.1110	0.1445	0.1620
$P(mm)$	0.1110	0.1498	0.1705
$P(md)$	0.1110	0.1509	0.1484
$P(du)$	0.1110	0.1455	0.1178
$P(dm)$	0.1110	0.1178	0.0907
$P(dd)$	0.1110	0.0594	0.0190
$P(u)$	0.3330	0.2321	0.2915
$P(m)$	0.3330	0.4453	0.4809
$P(d)$	0.3330	0.3226	0.2276

Table 4.3: Probability distributions on Ω'

probability distribution can be determined by P_3 and so on. Actually minimization through $\mathcal{P}$ determines the combination that minimizes the expected value of the portfolio X.

The last step is to find the risk adjusted value at t=0. Using the recursivity of ϕ, single period probability measures giving the minimum expected value at each non-final node will be determined. For instance, to find the minimum expected value at node uu;

$$\begin{aligned} & P_1(u \mid uu) \times X_3(uuu) + P_1(m \mid uu) \times X_3(uum) + P_1(d \mid uu) \times X_3(uud) \\ = \ & 0.333 \times 23634,75 + 0.333 \times 23516,34 + 0.333 \times 23472,44 \\ = \ & 23517.64 \end{aligned}$$

$$\begin{aligned} & P_2(u \mid uu) \times X_3(uuu) + P_2(m \mid uu) \times X_3(uum) + P_2(d \mid uu) \times X_3(uud) \\ = \ & 0.114 \times 23634,75 + 0.328 \times 23516,34 + 0.558 \times 23472,44 \\ = \ & 23505.23 \end{aligned}$$

$$
\begin{aligned}
& P_3(u \mid uu) \times X_3(uuu) + P_3(m \mid uu) \times X_3(uum) + P_3(d \mid uu) \times X_3(uud) \\
= \; & 0.161 \times 23634,75 + 0.2232 \times 23516,34 + 0.6158 \times 23472,44 \\
= \; & 23508.37
\end{aligned}
$$

The normal distribution (P_2) gives the minimum expected value conditional to the node uu. When this process is repeated for the other non final nodes, we end up with

$$\phi(X) = 22847.9$$

The product probability distribution for the given Ω is {0.00323,0.0093,0.0158,0.02855,0.02864,0.033,0.03514,0.03912,0.0395,0.0494, 0.0494,0.0494,0.0494,0.0494,0.0494,0.0494,0.0494,0.0494,0.3226,0.3226,0.3226, 0.3226,0.3226,0.3226,0.3226,0.3226,0.3226} respectively. It's composition is also shown in the Figure 4.3.

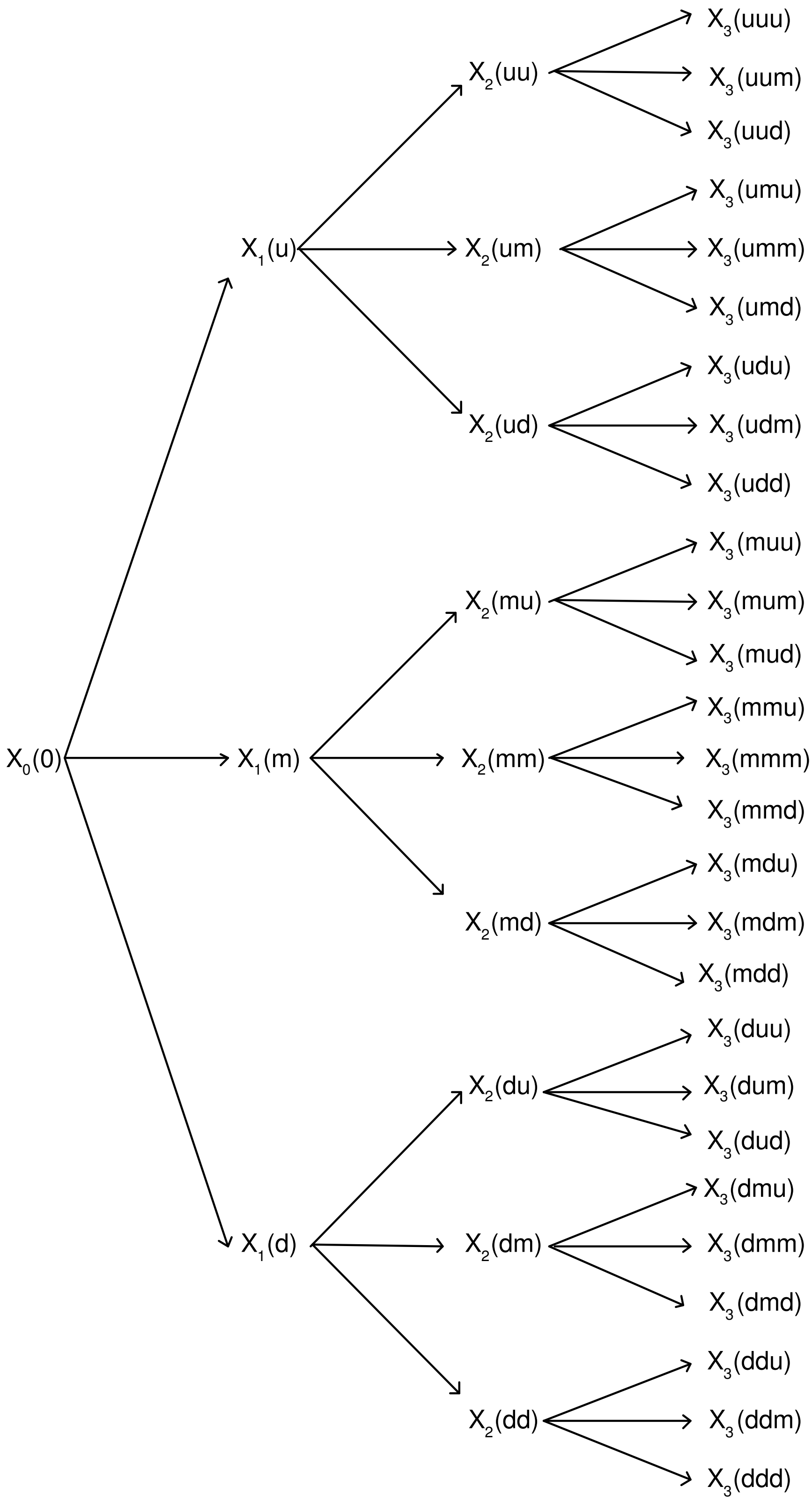

Figure 4.1: Value Process X

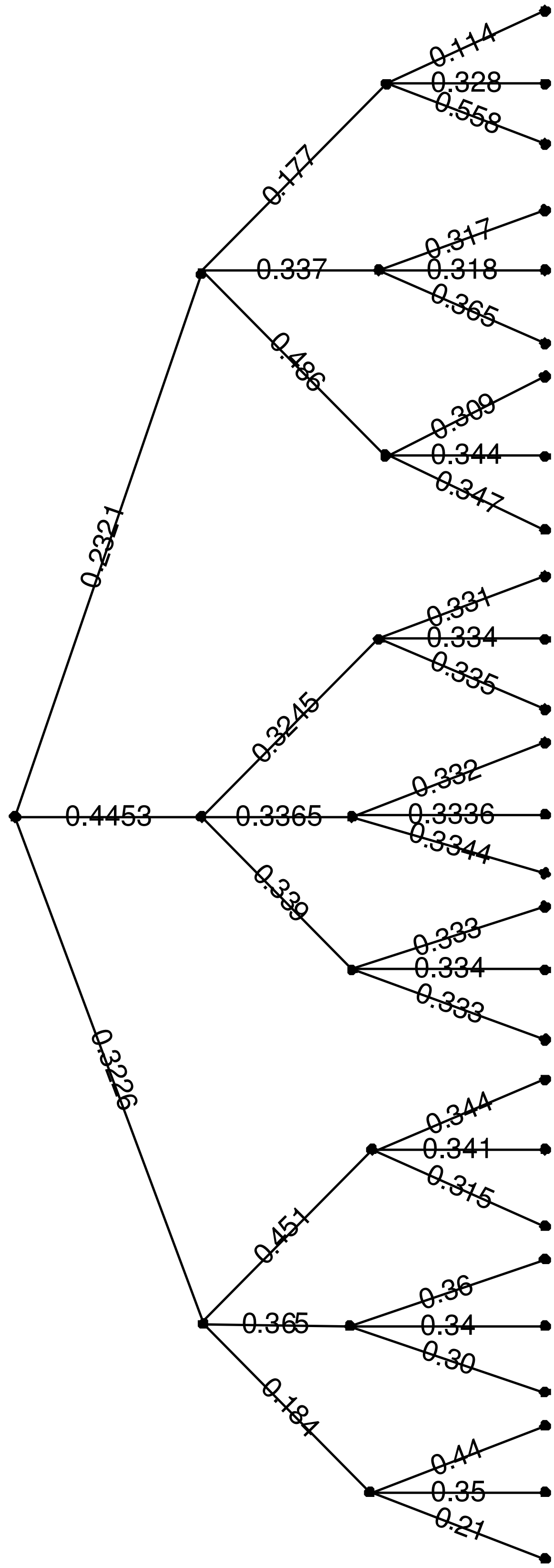

Figure 4.2: Single Period Probabilities of P_2

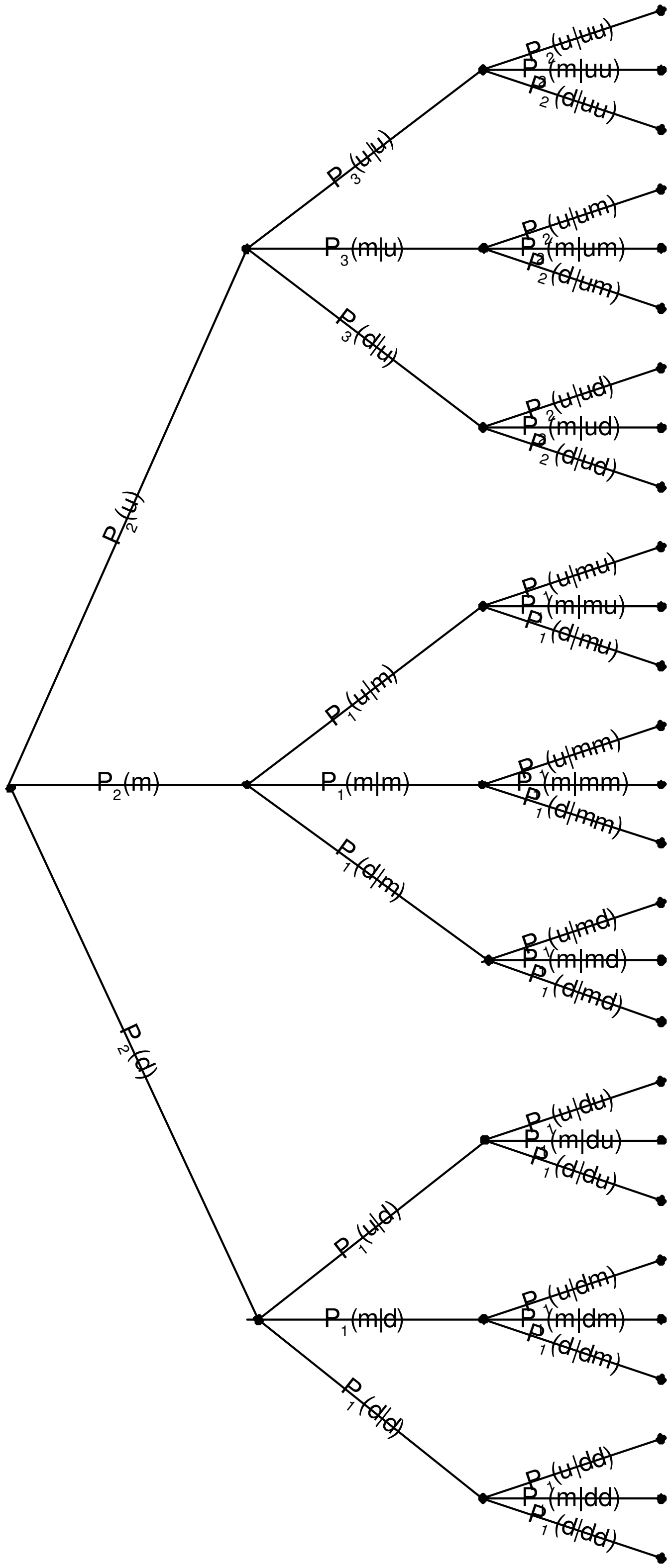

Figure 4.3: Product Probability Distribution

Chapter 5

CONCLUSION

In this work we reviewed risk measures, trying to model the market risk and determine the capital level required to cover possible losses. Since VaR does not encourage portfolio diversification, coherent risk measures were developed. Afterwards convex and conditional convex risk measures followed. Although these have stronger mathematical foundations than VaR, they can be a little complicated in application. Different types of consistent risk measures, easy to apply and interpret, are needed. Some alternatives like expected shortfall, conditional value at risk and risk measures using distorted probability measures are examined to handle this problem. Moreover, since financial decisions are not a one period process, multiperiod extensions are more realistic. Only dynamic models only for coherent risk measures are presented in this work.

As a further real data applications of convex risk measures would be considered. Furthermore, penalty functions of the convex risk measures can be formulated using the disutility function l. Such a formulation takes into account the level of risk aversion of an agent and combines the real world probabilities with subjective utility functions. Another subject can be investigating the conditional convex risk measures when the intermediate action is a possibility. When the insolvency time is estimated, best action, like an option contract or an insurance agreement can be made. The extension of multiperiod dynamic models to infinite set of states of world is another necessity for consistent risk measures to be widely accepted. Also the integration of the asymmetric information theory into the multiperiod setting is a promising area of study.

APPENDIX

A1.Finitely Additive and σ Finite Measures: Let $(\Omega, \mathcal{F})$ be a measurable space. A mapping $\mu : \mathcal{F} \to \mathbb{R}$ is called a finitely additive set function if $\mu(\emptyset) = 0$ and if it satisfies the following.

Whenever A_i are disjoint, $A_i \in \mathcal{F}$ for i=1,...,n and for

$$A = \bigcup_{i=1}^{n} A_i \in \mathcal{F}'$$

;

$$\mu(A) = \sum_{i=1}^{n} \mu(A_i).$$

$\mathcal{M}_{1,f}(\Omega, \mathcal{F})$ denotes the set of finitely additive set functions $\mu : \mathcal{F} \to [0, 1]$ which are normalized to $\mu(\Omega) = 1$. The total variation of a finitely additive set function μ is defined as

$$\|\mu\|_{var} := \sup\{\sum_{i=1}^{n} |\mu(A_i)| \mid A_i \textit{ are disjoint sets in } \mathcal{F}\}.$$

The space of all finitely additive measures μ whose total variation is finite is denoted $\mathbf{ba}(\Omega, \mathcal{F})$. Furthermore the space $\mathcal{X}$ of all bounded measurable functions on $(\Omega, \mathcal{F})$ is a Banach space if endowed with the supremum norm. $\mathcal{M}_{1,f}(\Omega, \mathcal{F})$ is contained in $\mathbf{ba}(\Omega, \mathcal{F})$ denotes the integral of a function $F \in \mathcal{X}$ with respect to $Q \in \mathcal{M}_{1,f}(\Omega, \mathcal{F})$ by

$$E_Q[F] = \int F dQ.$$

A measure μ is said to σ finite, when for all $F \in \mathcal{F}$; $\mu(F) < \infty$.

A2.Duality: If X is a topological vector space over R, all continuous linear mappings from X into $\mathbb{R}$ constitute the dual of X and are denoted by X'. That is, for every $x \in X$ and $f \in X'$; $f(x) \in \mathbb{R}$. Moreover, the value of the linear form f at point x is denoted as $< x, f >$.

A3.Weak and Weak* Topology: Let (X, X') is a dual pair. For any $f \in X'$ the functional defined by

$$p_f(x) = |f(x)| = |< x, f >| \; for \; all \; x \in X$$

is a seminorm on X. The coarsest topology on X, making all those seminorms continuous, is the weak topology and is denoted by $\sigma(X, X')$.

Dually, for any $x \in X$, we define a seminorm q_x on X' by

$$q_x(f) = |f(x)| \; for \; all \; f \in X'.$$

The coarsest topology on X', making all these seminorms continuous, is the weak* topology and is denoted by $\sigma(X', X)$.

Remark: Although mathematical definitions of the dual space, weak and weak* topologies are given above, they need some further interpretation. Every normed vector space X is, by using the norm to measure distances, a metric space and hence a topological space. This topology on X is also called the strong topology. The weak topology on X is defined using the continuous dual space X'. This dual space consists of all linear functions from X into $\mathbb{R}$ which are continuous with respect to the strong topology. The weak topology on X is the weakest topology (the topology with the fewest open sets) such that all elements of X' remain continuous. Explicitly, a subset of X is open in the weak topology if and only if it can be written as a union of (possibly infinitely many) sets, each of which being an intersection of finitely many sets of the form $f^{-1}(U)$ with f in X* and U an open subset of $\mathbb{R}$. Moreover weak topologies are convex.

The dual space X' is a normed vector space by using the norm $||f|| = sup_{||x|| \leq 1} |f(x)|$. This norm gives rise to the strong topology on X'. Weak* topology on X' is the weakest topology such that for every $x \in X$, the substitution map $\phi_x : X' \to \mathbb{R}$, defined as $\phi_x(f) = f(x)\ for\ all\ f \in X'$, remains continuous. Subsets of X', which are closed for $\sigma(X, X')$, are called weak* closed.

A4. Polar Sets: Let (X, X') be a dual pair. If C is a subset of X, the polar of C is denoted by C^o and is the subset of X' satisfying the properties given below.

1. C^o is absolutely convex and $\sigma(X', X)$ closed.
2. If $C \subset B$, then $B^o \subset C^o$.
3. If $\lambda \neq 0$, then $(\lambda A)^o = (\frac{1}{|\lambda|} A^o)$.
4. $(\bigcup_\alpha A_\alpha)^o = \bigcap_\alpha A_\alpha^o$.

A5. Bipolar Theorem: Let $< X, X' >$ be a duality. For any subset $C \subset X$, the bipolar M^{oo} is the $\sigma(X, X')$ closed convex hull of $C \cup \{0\}$. (That is the smallest convex set containing $C \cup \{0\}$.)

A6. Separation Theorem: In a topological vector space X, any two disjoint convex sets B and C, one of which has an interior point, can be separated by a non-zero continuous linear functional l on X, i.e.

$$l(x) \leq l(y)\ for\ all\ x \in C\ and\ all\ y \in B.$$

A7. The integral

$$l(F) = \int F d\mu,\ F \in \mathcal{X},$$

defines a one-to-one correspondence between continuous linear functionals l on $\mathcal{X}$ and finitely additive set functions $\mu \in \mathbf{ba}$.

A8. Relative Entropy: Relative entropy of Q with respect to P is

$$\varepsilon(Q, P) = E^Q[ln\frac{dQ}{dP}] = E^P[\frac{dQ}{dP} ln\frac{dQ}{dP}]$$

where P and Q are two equivalent probability measures (their null sets are same) on $(\Omega, \mathcal{F})$. If a strictly convex function $f(x) = xlnx$ is introduced,

$$\varepsilon(Q, P) = E^P[f(\frac{dQ}{dP})].$$

Relative entropy is a convex functional of Q. Due to Jensen's inequality $\varepsilon(Q, P) \geq 0$ with $\varepsilon(Q, P) = 0$ if and only if $\frac{dQ}{dP} = 1$ $a.s.$ Shortly, relative entropy measures the difference between two probability measures.

A.9 Single Period Probability Measures for P_1 and P_3:

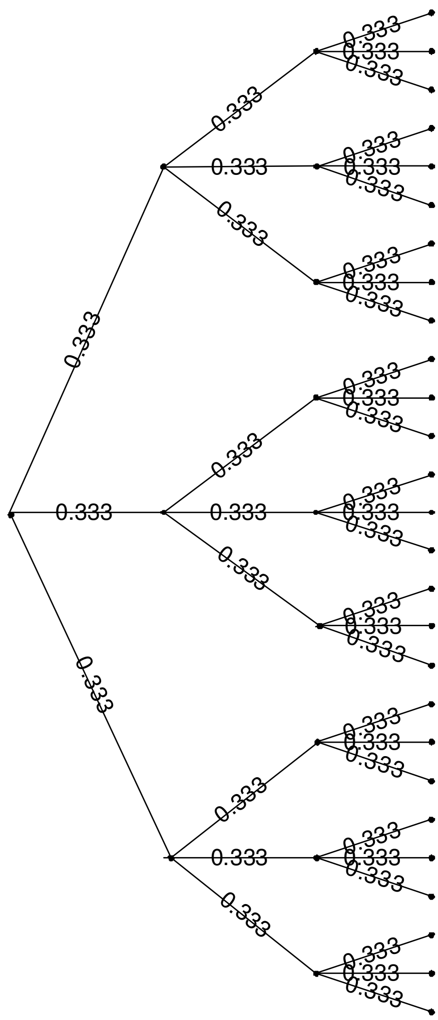

Figure 1: Single period probability measures for P_1

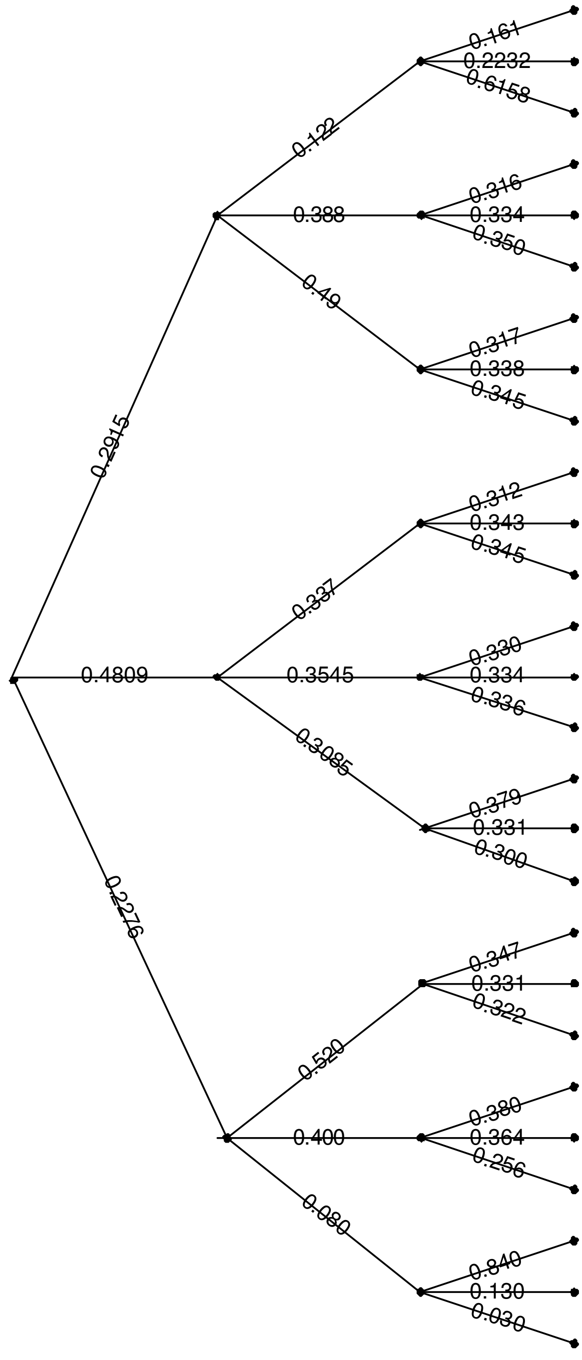

Figure 2: Single period probability measures for P_3

References

[AcNSi01] Acerbi C., Nordio C. and Sirtori C., *Expected Shortfall as a tool for Financial Risk Management*, Working Paper, AbaXBank (2001).

[AcT02a] Acerbi C. and Tasche D., *Expected Shortfall: a Natural Coherent Alternative to Value-at-Risk* , Economic Notes of Banca Monte dei Paschi di Siena SpA., 31(2), 1-10 (2002).

[AcT02b] Acerbi C. and Tasche D., *On the Coherence of Expected Shortfall*, Journal of Banking and Finance, 26(7) (2002).

[ADEH97] Artzner Ph., Delbean F., Eber J-M. and Heath D., *Thinking Coherently*, RISK 10, November, 68-71 (1997).

[ADEH99] Artzner Ph., Delbean F., Eber J-M. and Heath D., *Coherent Measures of Risk*, Mathematical Finance, 9, 203-228 (1999)

[ADEHKu02] Artzner Ph., Delbean F., Eber J-M., Heath D. and Ku H., *Multiperiod Risk and Coherent Multiperiod Risk Measurement*, Mimeo, ETH Zürich (2002).

[ADEHKu03] Artzner Ph., Delbean F., Eber J-M., Heath D. and Ku H., *Coherent Multiperiod Risk Adjusted Values and Bellman's Principle*, Manuscript, Université Louis Pasteur (2003).

[BN04] Bion-Nadal J., *Conditional Risk Measure and Robust Representation of Convex Cnditional Risk Measures*, Ecole Polytechnique, Centre de Mathématics Appliquées, R.I. No 557 (2004).

[D00] Delbean F., *Coherent Measures of Risk on General Probability Spaces*, Preprint, ETH Zürich (2000).

[D01] Delbean F., *Coherent Risk Measures*, Lecture Notes, Pisa (2001).

[DaM00] Danielsson J. and Morimoto J., *Forecasting Extreme Financial Risk: A Critical Analysis of Practical Methods for the Japanese Market*, Monetary Economic Studies, 12, 25-48 (2000).

[Do02] Dowd K., *An Introduction to Market Risk Measurement*, John Wiley & Sons (2002).

[DuSc58] Dunford N. and Schwartz J., *Linear Operators Vol 1*, Interscience Publishers, New York (1958).

[Dud89] Dudley R., *Real Analysis and Probability* , Wadsworth & Brooks/Cole, California (1989).

[E00] Embrechts P., *Extreme Value Theory: Potential and Limitations as an Integrated Risk Management Tool*, Working Paper (2000).

[E01] Embrechts P., *Extremes in Economics and Economics of Extremes*, SemStat meeting on Extreme Value Theory and Applications, Gothenburg (2001).

[Fi01] Fischer T., *Examples ofCoherent Risk Measures Depending on One-Sided Moments*, Working Paper, TU Darmstadt (2001).

[FrR02] Fritelli M. and Rosazza G., *Putting Order in Risk Measures*, Journal of Banking and Finance, 26, 1473-1486 (2002).

[FS02a] Föllmer H. and Schied A., *Convex Measures of Risk and Trading Constraints*, Finance and Stochastics, 6, 429-447 (2002).

[FS02b] Föllmer H. and Schied A., *Stochastic Finance: An Introduction in Discrete Time*, de Gruyter Studies in Mathematics, Berlin (2002).

[FS02c] Föllmer H. and Schied A., *Robust Representation of Convex Measures of Risk*, Advances in Finance and Stochastics, Essay in Honour of Dieter Sondermann, Springer-Verlag, New York (2002).

[G73] Grothendieck A., *Topological Vector Spaces*, Gordon and Breach, New York (1973).

[GeSe04] Gençay R. and Selçuk F., *Extreme Value Theory and Value-at-Risk: Relative Performance in Emerging Markets*, International Journal of Forecasting, 20, 287-303 (2004).

[GeSeUlu03] Gençay R., Selçuk F. and Ulugülyağci A., *High Volatility, Thick Tails and Extreme Value Theory in Value-at-Risk Estimation*, Insurance: Mathematics and Economics, 33, 337-356 (2003).

[Hol03] Holton G., *Value-at-Risk: Theory and Practice*, Academic Press (2003).

[Hu81] Huber P., *Robust Statistics*, Wiley, New York (1981).

[JP02] Jarrow R. and Purnanandam A., *Generalized Coherent Risk Measures: The Firm's Perspective* (2002).

[K04] Körezlioğlu H., *on Financial Risk Measures (A Quick Survey)*, Working Notes (2004).

[KHa01] Körezlioğlu H. and Hayfavi A., *Elements of Probability Theory*, METU, (2001).

[P00] Pflug G., *Some Remarks on the Value-at-Risk and the Conditional Value-at-Risk in Uryasev, S.*, Probabilistic Constrained Optimization: Methodology and Application, Kluwer Academic Pub. (2000).

[Pe03] Peter A., *Risk Measures*, SonderForschungsBereich 504, No 03-01 (2003).

[Rie04] Riedel F., *Dynamic Coherent Risk Measures*, Stoch. Proc. Appl., 112(2), 185-200 (2004).

[Ro70] Rockafeller R. T., *Convex Analysis*, Princeton University Press, New Jersey (1970)

[RoU00] Rockafeller R. T. and Uryasev S., *Optimization of Conditional Value-at-Risk*, The Journal of Risk, 2(3), 21-41 (2000).

[RoU01] Rockafeller R. T. and Uryasev S., *Conditional Value-at-Risk for General Loss Distributions*, Research Report 2001-5, ISE Dept. University of Florida (2001).

[RR73] Robertson A. P. and Robertson W., *Topological Vector Spaces*, Cambridge Univ. Press, (1973).

[RSE] Roorda B., Schumacher H. and Engwerda J., *Coherent Acceptibility Measures in Multiperiod Models*, Working Paper (2004).

[Sch71] Schaefer H., *Topological Vector Spaces*, Grad. Texts in Math. 3, Springer-Verlag, New York (1971).

[W02] Wang S. S., *A Risk Measure That Goes Beyond Coherence*, 12th AFIR International Colloquim, Mexico (2002).

[We02] West G., *Coherent VaR-Type Measures*, Financial Modelling Agency (2002).