Université de Montréal

On Some Aspects of Coherent Risk Measures and their Applications

par

Hirbod Assa

Département de mathématiques et de statistique

Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures

en vue de l'obtention du grade de

Philosophiæ Doctor (Ph.D.)
en Mathématiques

août 2011

Université de Montréal

Faculté des études supérieures

Cette thèse intitulée

On Some Aspects of Coherent Risk Measures and their Applications

présentée par

Hirbod Assa

a été évaluée par un jury composé des personnes suivantes :

Anatole Joffe

(président-rapporteur)

Bruno Rémillard, Manuel Morales

(directeur de recherche)

Jean-François Renaud

(membre du jury)

Benoit Perron

(représentant du doyen de la FES)

Thèse acceptée le:

juillet 2011

RÉSUMÉ

Le sujet principal de cette thèse porte sur les mesures de risque. L'objectif général est d'investiguer certains aspects des mesures de risque dans les applications financières. Le cadre théorique de ce travail est celui des mesures cohérentes de risque telle que définie dans [5]. Mais ce n'est pas la seule classe de mesure du risque que nous étudions. Par exemple, nous étudions aussi quelques aspects des "statistiques naturelles de risque" (en anglais *natural risk statistics*) [53] et des mesures convexes du risque [42]. Les contributions principales de cette thèse peuvent être regroupées selon trois axes: allocation de capital, évaluation des risques et capital requis et solvabilité. Dans le chapitre 2 nous caractérisons les mesures de risque avec la propriété de *Lebesgue* sur l'ensemble des processus bornés càdlàg (continu à droite, limité à gauche). Cette caractérisation nous permet de présenter deux applications dans l'évaluation des risques et l'allocation de capital. Dans le chapitre 3, nous étendons la notion de statistiques naturelles de risque à l'espace des suites infinies. Cette généralisation nous permet de construire de façon cohérente des mesures de risque pour des bases de données de n'importe quelle taille. Dans le chapitre 4, nous discutons le concept de "bonnes affaires" (en anglais *Good Deals*), pour notamment caractériser les situations du marché où ces positions pathologiques sont présentes. Finalement, dans le chapitre 5, nous essayons de relier les trois chapitres en étendant la définition de "bonnes affaires" dans un cadre plus large qui comprendrait les mesures de risque analysées dans les chapitres 2 et 3.

Mots-clés: mesures cohérentes et convexes de risque, propriété de Lebesgue, processus càdlàg, allocation de capital, statistiques naturelles de risque, couverture et tarification, bonnes affaires, capital requis et solvabilité.

SUMMARY

The aim of this thesis is to study several aspects of risk measures particularly in the context of financial applications. The primary framework that we use is that of coherent risk measures as defined in [5]. But this is not the only class of risk measures that we study here. We also investigate the concepts of natural risk statistics [53] and convex risk measure [42]. The main contributions of this Thesis can be classified in three main axes: Capital allocation, risk measurement and capital requirement and solvency. In chapter 2, we characterize risk measures with the *Lebesgue* property on bounded càdlàg processes. This allows to present two applications in risk assessment and capital allocation. In chapter 3, we extend the concept of *natural risk statistics* to the space of infinite sequences. This has been done in order to introduce a consistent way of constructing risk measures for data bases of any size. In chapter 4, we discuss the concept of Good Deals and how to deal with a situation where these pathological positions are present in the market. Finally, in chapter 5, we try to relate all three chapters by extending the definition of Good Deals to a larger set of risk measures that somehow includes the discussions in chapters 2 and 3.

Keywords: Coherent and Convex Risk Measure, Lebesgue Property, Càdlàg Process, Capital Allocation, Natural Risk Statistics, Hedging and Pricing, Good Deal, Capital Requirement, Solvency

CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENT

This research project would not have been possible without the support of many people. The author wishes to express his gratitude to his supervisor Prof. Bruno Rémillar and co-supervisor Prof. Manuel Morales, who were abundantly helpful and offered invaluable assistance, support and guidance. Deepest gratitude are also due to the member of the supervisory committee, Prof. Alejandro Bálbas without whose knowledge and assistance this study would not have been successful. Special thanks to , Prof. Anatole Joffe for his invaluable assistance and support during my studies at university of Montreal. The author would like to use this occasion to express his thanks to his late friend Prof. Isidore Fleischer with whose helpful comments, this thesis has a decent shape.

The author would also like to convey thanks to the Faculty, *the Institute of Mathematical Science, the Institute of Mathematical Finacne of Montreal (IFM2)* and *Le Fonds de recherche du Québec Nature et technologies (FQRNT)* for providing the financial means.

The author wishes to express his love and gratitude to his beloved families, in particular his wife (Fatemeh) and his mother; for their understanding and endless love, through the duration of his studies.

INTRODUCTION

Assessing financial risks is an ever present concern in economics and mathematical finance. The mathematical framework that incorporates a quantifiable financial risk was originally defined in terms of the language of von Neumann-Morgenstern expected utility theory, i.e., at the individual level, risk has always been characterized in terms of preference relations. Yet, from a risk management perspective, profits or losses are what define and quantify risk. The groundbreaking work of [5] is the cornerstone of a sound mathematical theory of risk measures that is compatible with risk management applications. They introduce the notion of risk measure as a real-valued function that assigns a meaningful numerical value to any given financial model. Their construction is axiomatic and it allows for a rich mathematical theory with room for practical applications. In fact, many examples of axiomatic risk measures are readily applied in practice and appear naturally in mathematical finance. Nowadays, risk measures have found their place as a relevant field in financial mathematics. The theory of risk measures is built with tools from well-developed fields of mathematics like probability and convex analysis. One element behind this success is that, although the axiomatic construction of these objects is dictated by the mathematical tools behind the theory, these also respond to financial intuition and needs. These axiomatic risk measures have mathematical representations that, far from being mere artifacts, have economical meaning. This brings new insight into the discussion.

The quality and volume of literature published about risk measures bears witness to the theoretical and practical interest that the seminal paper of [5] produced. A large amount of research followed, studying different aspects, implications and applications of the theory of coherent risk measures. As we will see, practical applications call for generalizations of the theory that will include a wide range of models. We find for instance works on risk measures defined on different spaces accordingly to particular needs. In [21, 22], the authors work out risk measures on the space of random processes modeling the outcome of a certain financial position; in [23] they develop risk measures in a dynamic fashion; in [49] they consider a set-valued risk measure instead of only real single valued (see also [46]). In [16], the authors extend the range of a coherent risk to a Banach space. We can also mention the work in [38] where they attempt to extend the risk measure on the largest possible space of all financial positions.

Nonetheless, there are several open questions and interesting directions yet to be explored. For instance, an argument can be made about the inadequacy of a simplistic solution measuring risk by means of a single real number [46]. In practice, risk managers desperately seek for simple positive or negative answers that can be easily decoded from risk measures (see a nice discussion in [31]). At a conceptual level, axiomatic risk measures do not have the risk-aversion feature that one would expect to see in any model that describes individual choices. The observable economical fact that individuals are generally averse to risk is not a part of the mathematical theory of risk measures [37].

This thesis explores some of these issues and produces new generalizations that seek to fill in gaps in the existing body of the theory and practice of risk measures. All of these extensions are not trivial since they call for the mathematical construction of suitable topologies. The study and extension of risk measures is a mathematical subject of interest in its own right. But the same can be said about the application aspects of the theory. Indeed, there is a large amount of literature dealing with a wide range of problems arising from applications of risk measures. In terms of applications, we focus in this thesis on the problems of capital allocation, data-based risk measuring and pricing and hedging of financial positions.

This thesis is then a compilation of three independent research articles that deal with different aspects, both theoretical and practical, of coherent and convex risk measures. Each one of these articles is presented in a single chapter. The main contributions of this thesis are contained in chapters 2, 3 and 4. An introductory chapter 1 is included to give a brief summary of the main definitions and results of the theory of risk measures as well as to lay down the main mathematical concepts and tools that are needed throughout the thesis. In the final chapter, chapter 5, we attempt to conclude by discussing several directions in which the work of this thesis can be extended.

We now give a brief account of the content of each of the main chapters.

Chapter 2 is based on the paper [7] entitled *Lebesgue Property of Risk Measures for Bounded Càdlàg Processes and Applications* and it deals mainly with the so-called *Lebesgue property*. The Lebesgue property is a continuity property which has been studied for coherent risk measures when the value of financial position is modeled with a single random variable. Here, we characterize this property for a risk measure on the space of bounded càdlàg processes, in several equivalent ways. Among them it is worthwhile to mention the equivalence between the Lebesgue property of a risk measure and the Lebesgue property of the associated static risk measure. An immediate application of our discussion is to approximate the risk of a random process with the risk of its time discretization approximation.

As a second application, we solve the problem of capital allocation via a gradient allocation approach. The problem of capital allocation has been the object of recent research (see for instance [35], [45], [18], [64] and [32]). In recent years, this problem has been analyzed with the tools given by the theory of risk measures. In fact, the problem of finding the risk contribution of each department in the overall company risk always involves an optimization procedure. This requires a certain notion of derivative for a risk measure. For instance, either using the concept of *risk contribution* ([26], [40] and [63]) or using the Euler Lemma on a positive homogeneous function ([55]) or using sub-gradient of a coherent risk measure ([33]), we always need to have a notion of the derivative of a risk measure in order to implement such optimizations.

In our more general setting, the value of a financial position is modeled by a bounded càdlàg process. We find a fair allocation when we deal with a general coherent risk measure. In particular, we pay more attention to some examples from finance. As a significant application, we find the exact formula for allocating the risky capital, when the surplus of an insurance company is modeled with a joint α-stable random process and the cumulative risk measure is used in order to estimate the required capital.

In Chapter 3, which is based on the paper [10] entitled *Risk Measures on the Space of Infinite Sequences* , we deal with definitions of suitable risk measures. In that chapter, we discuss the axiomatically definition of a data-based risk measure, the so-called *natural risk statistics*. This is a new type of risk measure defined in order to overcome some of the drawbacks associated with the sub-additivity feature of a coherent risk measure. In fact, sub-additivity excludes the most popular risk measure in practice, Value at Risk, from the family of coherent risk measures.

A problem with natural risk statistics is that they are defined for a fixed number of data. This is not very convenient while working with unknown number of data entries. In [10] (*joint with Manuel Morales*), the concept of natural risk statistics is extended to the space of infinite sequences in order to construct a consistent family of risk measures for any dimension. In this paper, we define natural risk statistics on the space of infinite sequences and then we show how one can construct a family of risk measures for finite dimensional spaces of every dimension. The statistical robustness of this family is also studied. In fact, we propose a way to construct a consistent family of risk measures for all data sizes.

Chapter 4, which is based on the article [9] entitled *Good Deals and Compatible Extension of Risk and Pricing Rule: A Regulatory Treatment* , deals with the problem of calculating an appropriate level of capital reserve (capital requirement) for a financial institution such as a bank or an insurance company is an ever-present concern for regulators. In fact, there is a world-wide trend moving towards establishing technical directives for financial institutions that set out rules for calculation of their capital requirement. For instance, in the European Union, we find two agreements that set up standards on how to compute solvency levels that would render financial markets more stable: Basel II (for financial institutions) and Solvency II (for insurance institutions). Some of these rules make use of risk measures such as Value at Risk (VaR) or Expected Shortfall in order to compute capital requirement (one can consult the website of the Bank for International Settlements at http://www.bis.org/ for further information).

In [9], we discuss the problem of capital requirement and solvency in the light of pathological positions called Good Deals. We study how risk measure-based capital requirements levels can create pathological situations. Indeed, one problem in capital requirement assessment of a financial position is that it is done without taking the interaction with market short prices into account. This can produce positions called *Good Deals*. A good deal is a financial position that simultaneously produces no risk and has no cost. We also discuss the problem of pricing and hedging a financial position with what we call the No-Good-Deal pricing method. A significant observation is how the choice of a risk measure can produce some pathological and unacceptable positions in the market called Good Deals. In fact, given a fixed pricing rule, the existence of such positions depends on the choice of the risk measure. We pursue the question of how a given risk measure can be modified in order to rule out Good Deals from the market. The main focus of that article is to give a recovery procedure that would modify a given coherent risk measure in a market in order to remove Good Deals. This is done in the context of capital requirement assessment of a financial position in reserve.

Chapter 1

PRELIMINARIES

1.1. Measuring Economic Risk

The problem of measuring the financial risk associated with any given financial position is of uttermost importance in economics and finance. The ultimate goal behind any attempt at designing risk measures is to coherently define a rational preference order within a set of positions that will allow market agents to make decisions. In the last decade, a comprehensive theory of risk measures has been developed. In this first chapter, we introduce the mathematical notions that are needed throughout the thesis. We also give a brief account of the content of each chapter while placing them in the context of recent developments in the theory. In particular, we discuss a few applications of our results to well-known financial problems.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $L^0(\Omega, \mathcal{F}, \mathbb{P})$ denote the space of all random variables (i.e. all measurable functions) on this probability space.

A financial position X is an element of $L^0(\Omega, \mathcal{F}, \mathbb{P})$ modeling an uncertain payoff.

A risk measure is a function $\rho : L^0 \to \mathbb{R}$ (or sometimes $\rho : L^0 \to \mathbb{R} \cup \{+\infty\}$) which defines a preference order on L^0 allowing a decision-maker to choose between any two given positions.

The traditional approach to measuring financial risks in economics is given in terms of the theory of rational decision-maker preferences and expected utility. Classical references are [30] and [48] where the behavior of market agents is described in terms of preference relations and the theory of von Neumann-Morgenstern expected utility. In this thesis, we follow a more modern school of thought that uses the notion of coherent risk measures as developed in modern financial mathematics; see, e.g., [5].

Following [48], we introduce the concept of expected utility in an axiomatic way. We start by defining the concept of preference relation for a rational decision maker. Let B be a subset of L^0. A *rational decision maker preference relation* $\preceq$ is a binary relation over a *choice set* B if $\preceq$ fulfills the following conditions:

(1) Completeness. For every $X, Y \in B$ either $X \preceq Y$, $X \succeq Y$ or $X \sim Y$ holds (where $X \sim Y$ means $X \preceq Y$ and $X \succeq Y$ hold simultaneously).

(2) Transitivity. For every $X, Y, Z \in B$ such that $X \succeq Y$ and $Y \succeq Z$, we have $X \succeq Z$.

(3) Independence. Let $X, Y \in B$ be two positions such that $X \succeq Y$ and let $\lambda \in (0, 1]$. For any position $Z \in B$ we have $\lambda X + (1-\lambda)Z \succeq \lambda Y + (1-\lambda)Z$.

(4) Continuity. Let $X, Y, Z \in B$ be three positions such that $X \succeq Y \succeq Z$. Then, there exists a $\lambda \in [0, 1]$ such that $Y \sim \lambda X + (1-\lambda)Z$.

A fundamental result is the celebrated von Neumann-Morgenstern formulation of expected utility (see [48]).

Theorem 1.1.1. *A preference relation $\preceq$ in B satisfying axioms (1)-(4) can always be represented as follows*

$$X \preceq Y \iff \mathbb{E}[u(X)] \leq \mathbb{E}[u(Y)], \forall X, Y \in L^1 \tag{1.1.1}$$

for some increasing concave function $u : B \to \mathbb{R}$.

This result defines a risk measure (better said preference measure) as a function on B through an utility function as follows

$$\rho_u : X \mapsto -\mathbb{E}[u(X)] \ , \quad \forall X \in B \ .$$

These utility-based risk measures are not compatible with tools and notions recently developed in the field of theoretical financial mathematics. In this chapter we give a brief account of this modern theory of risk measures and we discuss some of the main differences with respect to the expected utility approach. Basic concepts from functional analysis and stochastic processes are presented.

1.2. Technical Preliminaries

In this section we introduce the definitions, theorems and propositions that we frequently use.

1.2.1. Dual Spaces

The following discussion is mostly taken from [2] and [44]. We start with the following definition.

Definition 1.2.1. *Let B be a vector space endowed with a topology. The space B is a Topological Vector Space (TVS) if the addition of vectors and the multiplication by a scalar are continuous.*

Let B and E be two TVS and suppose that $\langle \cdot, \cdot \rangle : B \times E \to \mathbb{R}$ is bilinear.

Definition 1.2.2. *The weak topology on B induced by E is the coarsest topology on B for which $x \mapsto \langle x, e \rangle$ is continuous, for any $e \in E$. We denote this topology by $\sigma(B, E)$.*

Let $(X_\lambda)_{\lambda \in \Lambda}$ be a net in B, where Λ is a directed set. The net $(X_\lambda)_{\lambda \in \Lambda}$ converges to X in $\sigma(B, E)$ if $\langle X_\lambda, e \rangle \underset{\lambda}{\to} \langle X, e \rangle$ for all $e \in E$. We denote this convergence with $X_\lambda \xrightarrow[\lambda]{\sigma(B,E)} X$. Similarly one can define $\sigma(E, B)$.

1.2.2. Banach Spaces

Let B be a linear space. A norm $\|\cdot\|$ on B is a function from B to $\mathbb{R}^+ = [0, \infty)$ such that

(1) $\forall X \in B\,,\ \|X\| = 0$ iff $X = 0$.

(2) $\forall X \in B\,,\ t \geq 0\,,\ \|tX\| = t\|X\|$

(3) $\forall X, Y \in B\,,\ \|X + Y\| \leq \|X\| + \|Y\|$.

The linear space B is a normed space if its topology is induced by metric $d(X, Y) = \|X - Y\|$.

Definition 1.2.3. *A normed space $(B, \|\cdot\|)$ is called a Banach space if it is complete.*

For any Banach space B (or $(B, \|\cdot\|)$) the space of all linear and continuous functions from B to $\mathbb{R}$ is called the dual space and is denoted by B^*. The linear space B^* is a Banach space with the following norm

$$\|f\| = \sup_{X \in B,\, \|X\| \leq 1} |f(X)|.$$

Definition 1.2.4. *Let B be a Banach space. For any $f \in B^*$ and $X \in B$, one defines the bilinear relation $\langle X, f \rangle = f(X)$.*

Let B be a Banach space and B^* its dual. The weak topology on B is the topology $\sigma(B, B^*)$. Also, the topology $\sigma(B^*, B)$ on B^* is called the weak-star topology.

Let C be a subset of a linear space B. The set C is convex if $\lambda X + (1-\lambda)Y \in C$, $\forall X, Y \in C$ and $\lambda \in [0,1]$.

Theorem 1.2.1. *Let B be a Banach space and let C be a convex subset of B. The set C is closed in the norm topology if and only if it is closed in the weak topology.*

Theorem 1.2.2. *Every bounded set in B^* is relatively compact (i.e., its closure is compact) with respect to the weak star topology.*

Given a Banach space B, the bi-dual space B^{**} (dual of dual) contains B by the following embedding

$$X \mapsto (X(V) = \langle X, V \rangle\,, \forall V \in B^*).$$

Definition 1.2.5. *A Banach space B is called reflexive if the previous embedding is an automorphism.*

Theorem 1.2.3. *Any bounded set in a reflexive Banach space B is relatively weak compact, i.e., relatively compact w.r.t the weak topology.*

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space (not necessarily a probability triple). Interesting examples of reflexive Banach spaces are the function spaces $L^p(\Omega)$, for $1 < p < \infty$, defined as the space of all measurable functions X on Ω such that $\|X\|_{L^p} = (\int_\Omega |X|^p d\mu)^{\frac{1}{p}}$ is finite. Indeed, the linear space $L^p(\Omega)$ is a Banach space equipped with the L^p norm.

A subset C of a linear space B is a cone if

$$\forall X, Y \in C\,, \lambda \geq 0\,,\ X + Y \in C \ \text{ and } \ \lambda X \in C.$$

1.2.3. Stochastic Processes

The following discussion is mostly taken from [34] and [51].

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space which contains all null sets (i.e is complete). The expectation $\mathbb{E}$ is a functional defined as $\mathbb{E}(X) = \int_\Omega X d\mathbb{P}$ on $L^1(\Omega)$.

A random process $X = X(t, \omega)$ on $[0, T]$ is a function from $\Omega \times [0, T]$ to $\mathbb{R}$ which is measurable with respect to the σ-field $\sigma(\mathcal{F} \times \mathcal{B})$, where $\mathcal{B}$ is the Borel sets on $[0, T]$. We say that two random processes X and Y are indistinguishable if the following set is a null set (of measure zero)

$$\{\omega | \exists t \in [0, T]\,,\, X_t(\omega) \neq Y_t(\omega)\}\,.$$

As $X \leq Y$ we mean that the following set is a null set

$$\{\omega | \exists t \in [0, T]\,,\, X_t(\omega) > Y_t(\omega)\}.$$

A càdlàg process X is a random process such that the set

$$\{\omega | t \mapsto X_t(\omega) \text{ is right continuous and left limited } \}$$

is of measure one.

Let $\{\mathcal{F}_t\}_{t\in[0,T]}$ be a family of increasing σ-fields contained in $\mathcal{F}$. We say that $\{\mathcal{F}_t\}_{t\in[0,T]}$ satisfies the usual conditions if $\mathcal{F}_0$ contains all null sets and $\{\mathcal{F}_t\}_{t\in[0,T]}$ is right continuous, i.e.

$$\forall s \in [0, T)\,,\, \cap_{t>s}\mathcal{F}_t = \mathcal{F}_s.$$

A random process X is $\{\mathcal{F}_t\}_{t\in[0,T]}$-adapted if for any $t \in [0, T]$, X_t is $\mathcal{F}_t$ measurable.

A stopping time τ is a nonnegative random variable so that the set $\{\tau \leq t\}$ is $\mathcal{F}_t$ measurable for any $t \in [0, T]$. For any stopping time τ and σ define the following intervals

$$[\tau, \sigma[= \{(t, \omega) \in [0, T] \times \Omega | t < \sigma(\omega), t \geq \tau(\omega)\},$$

$$]\tau, \sigma] = \{(t, \omega) \in [0, T] \times \Omega | t \leq \sigma(\omega), t > \tau(\omega)\}.$$

The other intervals are defined in a similar way. The σ-field generated by

$$\{[\tau, \sigma[\,|\tau, \sigma \text{ are a stopping time }\}$$

is denoted by $\mathcal{O}$ and is called the *optional* σ-field. The σ-field generated with

$$\{]\tau, \sigma]\,|\tau, \sigma \text{ are a stopping time }\}$$

is denoted with $\mathcal{P}$ and is called the *predictable* σ-field.

Definition 1.2.6. *A random process X is predictable if it is $\mathcal{P}$-measurable and X is optional if it is $\mathcal{O}$-measurable.*

The following result is a strong result which is called the *Optional Projection Theorem.* For a proof, see [34].

Theorem 1.2.4. *For any bounded and measurable random process X on $[0,T]\times\Omega$ (not necessarily adapted) there exists a unique optional random process Y for which for any stopping time τ we have*

$$Y_\tau = \mathbb{E}[X_\tau|\mathcal{F}_\tau].$$

We denote this unique random process Y by $\Pi^{\mathrm{op}}(X)$.

Definition 1.2.7. *A stopping time τ is called predictable if there exists a sequence $\{\tau_n\}_{n=1,2,...}$ of stopping times such that $\tau_n < \tau$ and $\tau_n \uparrow \tau$.*

The following result is another strong result which is called the *Predictable Projection Theorem.* For a proof, see [34].

Theorem 1.2.5. *For any bounded and measurable random process X on $[0,T]\times\Omega$ (not necessarily adapted) there exists a unique predictable random process Y for which for any predictable stopping time τ we have*

$$\mathbb{E}[Y_\tau] = \mathbb{E}[X_\tau].$$

We denote this random process Y by $\Pi^{\mathrm{pr}}(X)$.

1.3. Coherent Risk Measure

The seminal paper [5] gave a mathematically rigorous construction of a risk measure. Indeed, the authors introduced the concept of a coherent risk measure that implicitly defines a preference relation on a subset of L^0 representing uncertain payoff values of market financial positions.

Definition 1.3.1. *Let $\mathcal{K}$ be a convex cone in L^0 containing $\mathbb{R}$ ($\mathbb{R}$ as the space of constant functions). A function $\rho : \mathcal{K} \to \mathbb{R}$ is a coherent risk measure if ρ is*

(1) positive homogeneous, i.e. $\rho(\lambda X) = \lambda\rho(X)$, $\forall X \in \mathcal{K}$ and $\lambda \in (0, +\infty)$.

(2) sub-additive, i.e. $\rho(X + Y) \leq \rho(X) + \rho(Y)$, $\forall X, Y \in \mathcal{K}$.

(3) translation invariant, i.e. $\rho(X + m) = \rho(X) - m$, $\forall X \in \mathcal{K}$ and $m \in \mathbb{R}$.

(4) decreasing, i.e. $\rho(X) \leq \rho(Y)$, $\forall X, Y \in \mathcal{K}$ such that $X \geq Y$ almost surely.

If axioms (1) and (2) are replaced by

(2') convexity, i.e. $\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y)$, $\forall X, Y \in \mathcal{K}$ and $\lambda \in [0, 1]$,

the risk measure ρ is called a convex risk measure. Note that (1) *and* (2) *imply* $(2')$*, so a coherent risk measure is a convex risk emasure.*

This axiomatic definition is the cornerstone of a very rich theory that draws its building blocks from functional analysis and has interesting economic interpretation. The preference relation associated with each risk measure is defined viz.

$$X \preceq_\rho Y \Leftrightarrow \rho(X) \geq \rho(Y)\,,\ \forall X, Y \in \mathcal{K}. \tag{1.3.1}$$

It is a straightforward exercise to see that $\preceq_\rho$ is both transitive and complete. [1]

This preference relation (although not rational!) gives economic interpretation:

- In economic terms, axiom *(1)* simply states that increasing exposure to a risky position implies a proportional increase of the risk level. Interestingly enough, this axiom in the definition of a coherent risk measure produces a new preference relation that cannot be reproduced in the expected utility

[1] Moreover it is reflexive and therefore defines a total pre-order or weak ordering (see [4]).

approach (we will see this later). This preference relation is given by

$$X \preceq_\rho Y \Leftrightarrow \lambda X \preceq_\rho \lambda Y \,,\, \forall \lambda > 0. \tag{1.3.2}$$

- Axiom *(2)* reproduces the widely accepted notion that risk can be reduced with diversification. In fact, this feature is also found in the expected utility approach to risk due to the concavity of the utility function.
- Axiom *(3)* endows coherent risk measures with a cash-invariance feature. This property produces a preference relation different from those in the expected utility approach. In terms of preference relation the cash- invariance axiom is

$$X \preceq_\rho Y \Leftrightarrow X + m \preceq_\rho Y + m \,,\, \forall m \in \mathbb{R}. \tag{1.3.3}$$

  As we will see, this axiom plays an important role in applications, particularly in the capital allocation problem.
- Finally, axiom *(4)* simply states that if the payoff of a financial position is always larger than the payoff of second position then, the associated risk measure preserves this order. In terms of preference relation this axiom becomes

$$X \leq Y \;\Rightarrow\; X \preceq_\rho Y. \tag{1.3.4}$$

As for the alternative axiom *(2')*, this is less restrictive condition than *(1)* and *(2)*. Convex risk measures were first introduced and studied in [42]. Coherent and convex risk measures form two distinct families that have been extensively studied in the literature. These families constitute the modern approach to risk measuring as introduced in financial mathematics.

In the field of financial mathematics, there are models that describe the behavior of financial positions not only as static random variables but as dynamic stochastic processes. As a consequence, the notion of a risk measure has to be adapted in order to continue to serve its purpose. This means that the space of financial positions, $\mathcal{K}$, in Definition 1.3.1 can be redefined according to our modeling needs. For instance, it could be L^2, L^∞ or the space of bounded càdlàg processes $\mathcal{R}^\infty$. Other particular applications might call for more simple spaces, for instance if we want to define risk measures for data sets then $\mathbb{R}^n$ would be a suitable space to work with. In each of the following chapters, it will be clearly stated on what space we will be working as well as the difficulties and advantages of doing so.

1.4. Robust Representation of a Coherent Risk Measure

In this section we state the main definitions and representation theorems for coherent risk measures as well as the new concept of natural risk statistics.

Definition 1.4.1. *A coherent risk measure $\rho : L^\infty \to \mathbb{R}$ is said to have the Fatou property if for any bounded sequence X_n in L^∞ (i.e., $\exists c > 0, \|X_n\|_{L^\infty} < c, \forall n \in \mathbb{N}$) converging in probability to $X \in L^\infty$, we have*

$$\rho(X) \leq \liminf \rho(X_n). \tag{1.4.1}$$

If equality holds with lim *(instead of* lim inf*) in (1.4.1) then ρ is said to have the Lebesgue property.*

1.4.1. Fenchel-Moreau Type Representation

The main results in the theory of coherent risk measures are representation theorems characterizing the set of risk measures. In this first subsection, we give a brief account of such results from a convex analysis perspective.

In convex analysis it is shown that convex functions, under some moderate conditions ([36]), can be represented as a supremum of affine functions (Fenchel-Moreau representation theorem). In the theory of risk measures, this yields the following result from [33]. But first we need to make one point clear. In all discussions in this thesis we identify a subset of absolutely continuous measures $\mathcal{P}$ with the set of its Radon-Nikodym derivatives i.e. $\{f \in L^1_+(\Omega) | \exists \mathbb{Q} \in \mathcal{P}, f = \frac{d\mathbb{Q}}{d\mathbb{P}}\}$.

Theorem 1.4.1. *For a coherent risk measure $\rho : L^\infty \to \mathbb{R}$ the following are equivalent*

(1) ρ is a coherent risk measure with the Fatou property.

(2) There is a L^1-closed set of probability measures $\mathcal{P} \subseteq L^1$ such that

$$\rho(X) = \sup_{\mathbb{Q} \in \mathcal{P}} E^{\mathbb{Q}}[-X]. \tag{1.4.2}$$

(3) The set $\{X \in L^\infty | \rho(X) \leq 0\}$ is a weak star closed convex set.

(4) ρ is a coherent risk measure which is continuous from above i.e. for any bounded and decreasing sequence X_n converging to X, $\rho(X) = \lim \rho(X_n)$.

In the literature, we find another alternative way of characterizing coherent risk measures and the information they provide in terms of acceptance sets. For more details on these representations, we refer the reader to [33].

1.4.2. Natural Risk Statistics

Natural risk statistics, as defiend in [53], is an alternative to coherent risk measures. This type of risk measure is defined on $\mathbb{R}^n$, as the space of data with length n. Before moving on further, note that in the definition of *Natural risk statistics*, the argument of risk measure is "loss" instead of "profit" or "pay-off".

Definition 1.4.2. *A function $\rho : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a natural risk statistics if,*

(1)

$$\rho(\lambda X) = \lambda \rho(X),\ \forall\, X \in \mathbb{R}^n\,, \forall \lambda \geq 0.$$

(2)

$$\rho(X + c\mathbf{1}) = \rho(X) + c,\, \forall\, X \in \mathbb{R}^n, c \in \mathbb{R},$$

where $\mathbf{1} = (\underbrace{1, \ldots, 1}_{n-times})$.

(3)

$$\rho(X) \leq \rho(Y)\ , \forall X \leq Y\ ,$$

where this inequality must be understood component wise.

(4) For any $X = (x_1, \ldots, x_n), Y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ such that

$$(x_i - x_j)(y_i - y_j) \geq 0$$

for all $j \neq i$, then

$$\rho(x_1 + y_1, \ldots, x_n + y_n) \leq \rho(x_1, \ldots, x_n) + \rho(y_1, \ldots, y_n).$$

(5)

$$\rho(X) = \rho(X^{ij})\ ,$$

for all $X \in \mathbb{R}^n$ and all $i, j > 0$. Here the sequence X^{ij} is the element in $\mathbb{R}^n$ which is equal component wise to X except for the i-th and j-th component which are interchanged.

We discuss these risk measures in more detail in Chapter 3. At this stage, we simply point out that the main difference with coherent risk measures lies in axiom *(4)*. In the definition of natural risk statistics, the sub-additivity feature has been replaced by a slightly more restrictive one. As we will see later, that is all is needed to introduce the statistical concept of robustness into the discussion.

1.5. Comparison with the Expected Utility Approach

In this section we compare the classical approach using expected utility and the modern concept of coherent risk measure.

Let ρ be a coherent risk measure defined on a cone $\mathcal{K} \subseteq L^0$. We say that Y is preferred to X, denoted $X \preceq_\rho Y$, if $\rho(X) \geq \rho(Y)$ for $X, Y \in \mathcal{K}$. Alternatively, given an expected utility $U : \mathcal{K} \to \mathbb{R}$ through $U(X) = \mathbb{E}[u(X)]$, this function produces a preference relation as follows: $X \preceq_U Y$ if $U(X) \leq U(Y)$ for any $X, Y \in \mathcal{K}$.

It is a straightforward exercise to show that for any coherent risk measure ρ, the preference relation $\preceq_\rho$ is reflexive, complete, transitive and continuous. However, $\preceq_\rho$ is not an independent relation in general, i.e., if $X \preceq_\rho Y$ for all $t \in (0, 1]$ and Z we have that $tX + (1-t)Z \preceq_\rho tY + (1-t)Z$. That is important since independence is one of the most important properties of a preference relation based on an Expected Utility. Indeed, it can be shown that it is independent only if ρ is linear. In order to see that, note that because of the positive homogeneity of ρ, the independence feature of this preference relation reduces to $X + Z \preceq_\rho Y + Z$ for all $X \preceq_\rho Y$ and Z. That is to say, if $\rho(X) \leq \rho(Y)$, then $\rho(X+Z) \leq \rho(Y+Z)$ for all Z. Letting $Z = -X$, we get that $\rho(X) \leq \rho(Y)$ if and only if $\rho(Y - X) \geq 0$. In the same way, letting $Z = -Y$ we get that $\rho(X) \leq \rho(Y)$ if and only if $\rho(X - Y) \leq 0$. These two relations imply that $\rho(X) = \rho(Y)$ if and only if $\rho(X - Y) = \rho(Y - X) = 0$. Since $\rho(X - Y + \rho(X - Y)) = 0$, we have that $\rho(X) = \rho(Y) + \rho(X - Y)$. If we let now $Y = -X$, we have $\rho(X) = \rho(-X) + \rho(2X)$, which implies $\rho(X) = -\rho(X)$. Finally, by the Fenchel-Moreau representation (1.4.2) of ρ, that is possible only if $\rho(X) = \mathbb{E}^{\mathbb{Q}}[-X]$ for a given $\mathbb{Q}$.

This implies that the only function which is a coherent risk measure and for which the associated preference relation is that of a rational decision maker is $\rho : X \mapsto \mathbb{E}[-X]$. This is one of the major differences between these two approaches to risk.

In the approach to risk using expected utility, one key idea is the modeling of the risk averseness of individuals via a utility function $u : \mathbb{R} \to \mathbb{R}$. The function u is an increasing and concave function which shows the tendency to the larger payoffs while the rate of risk taking is restricted by the concavity of u. It turns out that if we want to have a risk measure-based rational preference, then the only choice for the utility function is the identity which corresponds to that of a risk neutral decision maker.

Under these considerations, it is clear that a coherent risk measure is not representing the risk averseness of individuals since it is equivalent to a risk-neutral utility function. However, by looking at the representation (1.4.2) of a coherent risk measure, one can see that what seems to matter to a decision maker is the uncertainty (ambiguity) surrounding the different scenarios in future events. This ambiguity is represented by the set of equivalent measures in the representation (1.4.2). This is what we call uncertainty or ambiguity aversion.

1.6. Some Aspects of Coherent Risk Measures

There are several financial problems that have been revisited with the concept of a coherent risk measure in the past few years. For instance, portfolio choice and asset allocation problems have been discussed in [60] and [54]. Applications of risk measures to the problem of capital allocation can be found in [35], while applications in optimal investment with convex risk measures are discussed in [61]. The problem of pricing and hedging in incomplete markets has also been the subject of interesting applications of risk measures like those in [41], [57], [62].

We find large amounts of theoretical and practical research that revolve around the definition and representation of coherent risk measures. Some generalizations have been necessary as different models for financial positions are required. Generalizing the theory of coherent risk measures to different spaces is an ongoing effort and at least two of the contributions of this thesis are in that direction. In Chapter 2, we extend some existing results for coherent risk measures defined on the space of càdlàg processes to convex risk measures. In this setting, financial position are modeled dynamically by a stochastic process for which a risk measure is needed. In Chapter 3, we explore the case of risk measures defined on the space of infinite data vectors l^∞. In that setting, risk positions are modeled by data sets available to a risk manager who must assess their associated risk. On the other hand, Chapter 4 discusses risk measures in the more classical space of $L^p(\Omega)$, $1 \leq p \leq \infty$, representing the payoff of a given financial position. The main contribution lies in the study and characterization of pathological positions called *Good Deals*.

Overall, this thesis is about some aspects of the applications of risk measures, namely,

(1) capital allocation,

(2) risk measurement,

(3) capital requirement and solvency.

Chapter 2

LEBESGUE PROPERTY OF RISK MEASURES FOR BOUNDED CÀDLÀG PROCESSES AND APPLICATIONS

Résumé

Dans cet article, nous étudions la propriété dite de *Lebesgue* pour des mesures convexes de risque sur un sous-ensemble de processus càdlàg. Nous généralisons les travaux de [32] et [50]. Pour cela, nous caractérisons les sous-ensembles compacts d'une famille de processus à variation bornée qui est le dual topologique des processus càdlàg, bien entendu, dans une topologie appropriée. Nous montrons que la propriété de *Lebesgue* peut être caractérisée de plusieurs façons équivalentes. Finalement, nous présentons des applications en évaluation de risque et en répartition de capital.

Abstract

In this paper, we study the so-called *Lebesgue* property for convex risk measures for a class of càdlàg processes. Our results extend previous work of [32] and [50]. We characterize the compact subsets of a family of the space of bounded variation processes which is the topological dual of the càdlàg processes, of course, in an appropriate topology. We show that the *Lebesgue* property can be characterized in several equivalent ways. Applications to risk assessment and allocation of risk

capital are presented.

2.1. Introduction

Coherent risk measures for finite probability spaces were introduced in [5] and were extended to general probability spaces in [33], where applications to risk measurement, premium calculation and capital allocation problems were discussed. In [42] the authors defined a more general notion of convex risk measures, and the representation results of [33] are extended. In [21, 22], the authors studied risk measures for stochastic processes, instead considering only random variables.

As can be seen in [32], the key concept for obtaining representations of a convex risk measure is the so-called *Fatou* property. This property is an order continuity for decreasing sequences in an appropriate space. The *Lebesgue* property is a stronger concept. In an appropriate space, it is related to a continuity property for uniformly bounded sequences, allowing for approximations of risk measures. It is also an order continuity for increasing sequences somehow completing its counterpart, the *Fatou* porperty. In the context of coherent risk measures for random variables, the *Lebesgue* property was studied in [33], while it was studied for convex risk measures on the space of random variables in [50].

In this paper we extend the definition of the *Lebesgue* property to the space of bounded càdlàg processes. We characterize the risk measures with *Lebesgue* property in several equivalent ways. Our main goal is to find equivalent conditions for the *Lebesgue* property in terms of conditions that can be readily verified. Having conditions that can be verified easily allows us to identify this property for complicated, but also interesting, convex risk measures. We consider two applications of our results in this paper. The first application follows directly from the definition of the *Lebesgue* property itself, which allows us to approximate a convex risk measure of a bounded càdlàg process X with the risk associated to an uniformly bounded sequence X_n converging to X. This is important when we deal with a time discretization of a finite time horizon. This type of approximation can be carried out using the *Fatou* property only if the approximating processes are a decreasing sequence which cannot always be carried out with time discretization. Now, having a *Lebesgue* property, we only need a uniformly bounded sequence of approximating processes. This is going to be discussed in Section 2.4 where we present a first round of examples. The second application is the use of the *Lebesgue* property in the capital allocation problem. In Theorems 2.5.2 and 2.5.3 we will show why the *Lebesgue* property is needed to have a fair allocation of risk capital. Then, we illustrate with a second round of examples how we can use our results to give an allocation.

It is also important to mention that one interesting contribution of this paper is the introduction of a Cumulative-Stopping risk measure. The use of such a measure is illustrated in an insurance application using an α-stable model.

The paper is organized as follows. In Section 2, we recall basic definitions and results for convex risk measures of random variables and for a class of càdlàg processes. In particular, we state two results, one related to the *Fatou* property for risk measures on the spaces of bounded càdlàg processes, and another one related to the *Lebesgue* property for risk measures on the space of bounded random variables. The theoretical results of the paper are presented in Section 3. In particular, we characterize relatively compact subsets of a given dual space and we characterize the *Lebesgue* property. Furthermore, we present an extended version of James' Theorem. In Section 4, we give some examples of risk measures with *Lebesgue* property. In Section 5 applications in capital allocation problem will be discussed. The proof of the theoretical results are given in the Appendix.

2.2. Preliminaries and Remarks

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a standard and atom-less probability space and let $(\mathcal{F}_t)_{0\leq t\leq T}$ be a filtration with the usual conditions. Furthermore, assume that $L^1(\Omega, \mathcal{F})$ has a countable dense subset.

In [21, 22] the authors developed the theory of convex risk measures on the space of $\mathcal{R}^p$ consisting of stochastic processes on $[0,T]$ that are càdlàg, adapted and such that $X^* = \sup_{[0,T]} |X_t| \in L^p$, with $1 \leq p \leq \infty$. Note that for any $1 \leq p \leq \infty$, $\mathcal{R}^p$, endowed with the norm $\|X\|_{\mathcal{R}^p} = \|X^*\|_{L^p}$, is a Banach space.

For $q \in [1, \infty]$, let $\mathcal{A}^q$ be the set of all $a = (a^{\mathrm{pr}}, a^{\mathrm{op}}) : [0,T] \times \Omega \to \mathbb{R}^2$ such that a^{pr} and a^{op} are right continuous, have finite variation in L^q, a^{pr} is predictable, and $a_0^{\mathrm{pr}} = 0$, a^{op} is optional and purely discontinuous.

Denoting the variation of a function $f : [0,T] \to \mathbb{R}$ by $\mathrm{Var}(f)$, it follows that $\mathcal{A}^q$ is also a Banach space, when equipped with the norm $\|a\|_{\mathcal{A}^q} = \|\mathrm{Var}(a)\|_{L^q}$. Furthermore, if p and q satisfy $\frac{1}{p} + \frac{1}{q} = 1$, there is a duality relation between $\mathcal{A}^q$ and $\mathcal{R}^p$,

$$\langle X, a \rangle = \mathbb{E}\left[\int_{]0,T]} X_{t-} da_t^{\mathrm{pr}} + \int_{[0,T]} X_t da_t^{\mathrm{op}}\right], \quad (X,a) \in \mathcal{R}^p \times \mathcal{A}^q. \tag{2.2.1}$$

Note that

$$|\langle X, a \rangle| \leq \|X\|_{\mathcal{R}^p} \|a\|_{\mathcal{A}^q}.$$

The subset $\mathcal{A}_+^q$ of $\mathcal{A}^q$ consisting of $a = (a^{pr}, a^{op})$ with both components non-negative and non-decreasing, will be important in the sequel.

Further, let $\mathcal{D}_\sigma$ be the unit ball of $\mathcal{A}_+^1$, i.e., the subset of $a \in \mathcal{A}_+^1$ such that

$$\|a\|_{\mathcal{A}^1} = E\left(a_T^{\mathrm{pr}} + a_T^{\mathrm{op}} - a_0^{\mathrm{op}}\right) = 1.$$

We are now in a position to recall some important definitions.

Definition 2.2.1. *A convex risk measure ρ on $\mathcal{R}^p$ is a function from $\mathcal{R}^p \to \mathbb{R}$ such that for any $X, W \in \mathcal{R}^p$:*

(1) $\rho(\lambda X + (1-\lambda)W) \leq \lambda\rho(X) + (1-\lambda)\rho(W)$, for all $0 \leq \lambda \leq 1$.

(2) $\rho(X + m) = \rho(X) - m$, for any $m \in \mathbb{R}$.

(3) $\rho(X) \geq \rho(W)$, whenever $X \leq W$.

ρ is called a coherent risk measure if in addition

(4) $\rho(\lambda X) = \lambda\rho(X)$ for any $\lambda > 0$.

In [21], the authors propose the following definition for the *Fatou* property for a convex risk measure on $\mathcal{R}^\infty$.

Definition 2.2.2. *A convex risk measure ρ on $\mathcal{R}^\infty$ has Fatou property if for any bounded sequence $\{X_n\}_{n \in \mathbb{N}} \subseteq \mathcal{R}^\infty$, for which there exists $X \in \mathcal{R}^\infty$ so that $(X_n - X)^* \xrightarrow{\mathbb{P}} 0$, we have $\rho(X) \leq \liminf \rho(X_n)$.*

The following characterization of the Fatou property for convex risk measures is taken from [21]. Recall that γ is a penalty function if $\gamma : \mathcal{D}_\sigma \to (-\infty, +\infty]$ is such that $-\infty < \inf_{a \in \mathcal{D}_\sigma} \gamma(a) < \infty$.

Theorem 2.2.1. *Let ρ be a mapping from $\mathcal{R}^\infty$ to $\mathbb{R}$. Then the following statements are equivalent.*

1-

$$\rho(X) = \sup_{a \in \mathcal{D}_\sigma} \left\{ \langle -X, a \rangle - \gamma(a) \right\}, \ X \in \mathcal{R}^\infty, \tag{2.2.2}$$

for some penalty function γ.

2- ρ is a convex risk measure on $\mathcal{R}^\infty$ such that $\left\{ X \in \mathcal{R}^\infty | \rho(X) \leq 0 \right\}$ is $\sigma(\mathcal{R}^\infty, \mathcal{A}^1)$-closed.

3- ρ is a convex risk measure on $\mathcal{R}^\infty$ with the Fatou property.

4- ρ is a convex risk measure on $\mathcal{R}^\infty$ which is continuous for bounded decreasing sequences.

Moreover, in each case, the conjugate function ρ^*, restricted to $\mathcal{D}_\sigma$, and defined by

$$\rho^*(a) = \sup_{X \in \mathcal{R}^\infty} \left\{ \langle -X, a \rangle - \rho(X) \right\},$$

is a penalty function which is smaller than γ and γ can be replaced by ρ^* in (2.2.2).

Remark 2.2.1. *As mentioned in [21], ρ^* restricted to $\mathcal{D}_\sigma$ equals $\rho^{\#}$ defined as follows:*

$$\rho^{\#}(X) := \sup_{a \in \mathcal{A}_\rho} \langle -X, a \rangle,$$

where $\mathcal{A}_\rho := \{ X \in \mathcal{R}^\infty \,|\, \rho(X) \leq 0 \}$ is the acceptance set of ρ. This implies that

$$\rho(X) = \sup_{a \in \mathcal{D}_\sigma} \left\{ \langle -X, a \rangle - \rho^{\#}(a) \right\}, \ X \in \mathcal{R}^\infty. \tag{2.2.3}$$

The following corollary is also taken from [21].

Corollary 2.2.1. *A coherent risk measure ρ on $\mathcal{R}^\infty$ has Fatou property if and only if there exists a subset $\mathcal{Q}$ of $\mathcal{D}_\sigma$ such that*

$$\rho(X) = \sup_{a \in \mathcal{Q}} \langle -X, a \rangle. \tag{2.2.4}$$

In general $\mathcal{Q}$ is not unique. An appropriate choice for $\mathcal{Q}$ is $dom(\rho^*) \cap \mathcal{A}^1_+ = \{a \in \mathcal{A}^1_+ \mid \rho^*(a) = 0\}$. In fact, due to positive homogeneity, one ends up with $\rho^*(a) = \lambda\rho^*(a)$ for any $\lambda > 0$, showing that $\rho^*(a) \in \{0, +\infty\}$.

Next, the *Lebesgue* property for risk measures on L^∞ was studied in [50], where the authors propose the following definition for *Lebesgue* property:

Definition 2.2.3. *A convex risk measure ρ on L^∞ has Lebesgue property if for any bounded sequence $\{Y_n\}_{n\in\mathbb{N}} \subseteq L^\infty$ converging to $Y \in L^\infty$ in probability, we have $\rho(Y) = \lim \rho(Y_n)$.*

Remark 2.2.2. *Notice that our definition is weaker because we use convergence in probability, whereas in [50] they use a.s. convergence in their definition of the Lebesgue property. We choose to use convergence in probability because this is important for our purpose. In fact, in [50] almost sure convergence could be replaced with convergence in probability. Indeed, all they need to derive their results is the fact that for any uniformly bounded sequence of random variables Y_n converging a.s. to Y and any uniformly integrable set C we have $\lim\limits_{n\to\infty} \inf\limits_{f\in C} \mathbb{E}[Y_n f] = \inf\limits_{f\in C} \mathbb{E}[Y f]$. The latter is also true if the convergence of Y_n is in probability instead of a.s. We refer to the proof of Theorem 3.6 in [32] for a thorough discussion.*

We extend the definition of the *Lebesgue* property to convex risk measures on $\mathcal{R}^\infty$ as follows.

Definition 2.2.4. *A convex risk measure ρ on $\mathcal{R}^\infty$ has Lebesgue property if for any bounded sequence $\{X_n\}_{n\in\mathbb{N}} \subseteq \mathcal{R}^\infty$, for which there exists $X \in \mathcal{R}^\infty$ so that $(X_n - X)^* \xrightarrow{\mathbb{P}} 0$, we have $\rho(X) = \lim \rho(X_n)$.*

Before giving the characterization theorem of convex risk measures with *Lebesgue* property we would like to recall that every convex risk measure ρ on L^∞ which has *Fatou* property can be represented as

$$\rho(Y) = \sup_{f\in L^1} \{\mathbb{E}[-fY] - \rho^*(f)\}. \tag{2.2.5}$$

By translation-invariance it turns out that

$$\rho(Y) = \sup_{f\in D_\sigma} \{\mathbb{E}[-fY] - \rho^*(f)\}, \tag{2.2.6}$$

where $D_\sigma := \{f \in {L^1}_+ \,|\, \mathbb{E}[f] = 1\}$ and ρ^* is the conjugate function on L^1. Also it is pointed out in [42] that

$$\rho^{\#}(f) := \sup_{Y \in \mathcal{A}_\rho} E[-fY] = \rho^*(f)\,,\ \forall f \in D_\sigma, \tag{2.2.7}$$

where $\mathcal{A}_\rho = \{Y \in L^\infty \,|\, \rho(Y) \le 0\}$.

The following result, proved in [50], is a characterization of convex risk measures on L^∞ with *Lebesgue* property.

Theorem 2.2.2. *Let ρ be a convex risk measure on L^∞ with Fatou property. The following conditions are equivalent.*

1- ρ has Lebesgue property.

2- $\left\{f \in L^1_+ | \rho^*(f) \le c\right\}$ is a $\sigma(L^1, L^\infty)$-compact subset of L^1 for every $c \in \mathbb{R}$.

3- $\mathrm{dom}(\rho^*) = \{\rho^* < \infty\} \subseteq L^1$.

4- In the representation (2.2.5) *the maximum is attained.*

Remark 2.2.3. *Following the same proof of Theorem 2.4 in [50] one can deduce that all expressions in the last theorem are equivalent to the following conditions.*

5- $\left\{f \in D_\sigma | \rho^{\#}(f) \le c\right\}$ is a $\sigma(L^1, L^\infty)$-compact subset of L^1 for every $c \in \mathbb{R}$.

6- In the representation (2.2.6) *the maximum is attained.*

We are now justified in extending the definition of the *Lebesgue* property to $\mathcal{R}^p$ for $1 \le p < \infty$ as follows.

Definition 2.2.5. *A convex risk measure ρ on $\mathcal{R}^p$, $1 \le p < \infty$, has Lebesgue property if the set $\{a \in \mathcal{A}^q : \rho^*(a) \le c\}$ is $\sigma(\mathcal{A}^q, \mathcal{R}^p)$-compact, where*

$$\rho^*(a) = \sup_{X \in \mathcal{R}^p} \{\langle X, a\rangle - \rho(X)\}\,, \quad a \in \mathcal{A}^q.$$

Remark 2.2.4. *We will see in the next section, Proposition 2.3.1, that as long as ρ has a representation like 2.2.2 (with $\mathcal{A}^q \cap \mathcal{D}_\sigma$ instead of $\mathcal{D}_\sigma$) then for the case $1 \le p < \infty$, ρ always has Lebesgue property. Theorem 2.3.2 shows that for the case $p = \infty$, ρ has Lebesgue property iff $\{a \in \mathcal{A}^1 : \rho^*(a) \le c\}$ is $\sigma(\mathcal{A}^1, \mathcal{R}^\infty)$-compact, which shows that definition 2.2.5 could also be extended for $p = \infty$.*

Before giving the theoretical results of the paper we should give some explanations and remarks which will be used in next discussions.

Define

$$\hat{\mathcal{R}}^p = \left\{ X : [0,T] \times \Omega \to \mathbb{R} \;\middle|\; \begin{array}{c} X \text{ is càdlàg} \\ X^* \in L^p \end{array} \right\}, \tag{2.2.8}$$

and

$$\hat{\mathcal{A}}^q = \left\{ a : [0,T] \times \Omega \to \mathbb{R}^2 \;\middle|\; \begin{array}{l} a = (a^{\mathrm{l}}, a^{\mathrm{r}}), a_0^l = 0 \\ a^{\mathrm{l}}, a^{\mathrm{r}} \text{ measurable} \\ \text{finite variation} \\ \text{and right continuous} \\ \mathrm{Var}(a^{\mathrm{l}}) + \mathrm{Var}(a^{\mathrm{r}}) \in L^q \end{array} \right\}. \tag{2.2.9}$$

Furthermore, extend the duality relation (2.2.1) by setting

$$\langle X, a \rangle = \mathbb{E}\left[\int_{]0,T]} X_{t-} da_t^{\mathrm{l}} + \int_{[0,T]} X_t da_t^{\mathrm{r}} \right], \quad (X,a) \in \hat{\mathcal{R}}^p \times \hat{\mathcal{A}}^q. \tag{2.2.10}$$

Remark 2.2.5. *By Theorems 65, 67 of section VII, [34], when* $p \neq \infty$*, the set* $\hat{\mathcal{A}}^q$ *is the dual of* $\hat{\mathcal{R}}^p$*. More precisely, when* $1 < p < \infty$*,* $\mathcal{A}^q$ *is the topological dual of* $\mathcal{R}^p$*, for any filtration* $(\mathcal{F}_t)_{t \in [0,T]}$*. For the case* $p = 1$ *and* $q = \infty$*, this happens only if* $\mathcal{F}_t$ *is constant and equals* $\mathcal{F}$ *for all* $t \in [0,T]$*. In general, for* $p = \infty$ *the equality* $(\mathcal{R}^\infty)^* = \mathcal{A}^1$ *does not hold, even if* $\mathcal{F}_t = \mathcal{F}$ *for all* $t \in [0,T]$*. This makes the case* $p = \infty$ *more difficult since it requires the use of more techniques and methods from functional analysis and the general theory of stochastic processes.*

Remark 2.2.6. *Denote with* Π^{op}, Π^{pr} *the optional and predictable projections as well as with* $\tilde{\Pi}^{op}$ *and* $\tilde{\Pi}^{pr}$ *the dual optional and predictable projections. See, e.g., [34], [51] or [21]. For* $a = (a^l, a^r) \in \hat{\mathcal{A}}^q$*, let* $\tilde{a}^l = \tilde{\Pi}^{pr}(a^l)$ *and* $\tilde{a}^r = \tilde{\Pi}^{op}(a^r)$*. One can split* $\tilde{a}^r$ *uniquely into a purely discontinuous finite variation part* $\tilde{a}^r_d$ *and a continuous finite variation part* $\tilde{a}^r_c$ *with* $\tilde{a}^r_c(0) = 0$*. Since* $\tilde{a}^r_c$ *is predictable, one can define a map* Π^* *from* $\hat{\mathcal{A}}^q$ *to* $\mathcal{A}^q$ *by*

$$\Pi^* a := (\tilde{a}^l + \tilde{a}^r_c, \tilde{a}^r_d).$$

Every predictable process is also optional, so $\tilde{a}^l, \tilde{a}^r_c, \tilde{a}^r_d$ *are all optional. It follows from [21] that*

$$\langle X, a \rangle = \langle X, \Pi^*(a) \rangle, \qquad (X,a) \in \mathcal{R}^p \times \hat{\mathcal{A}}^q. \tag{2.2.11}$$

Remark 2.2.7. (2.2.11) *implies that*

$$\Pi^* : (\hat{\mathcal{A}}^q, \sigma(\hat{\mathcal{A}}^q, \hat{\mathcal{R}}^p)) \to (\mathcal{A}^q, \sigma(\mathcal{A}^q, \mathcal{R}^p))$$

is continuous.

Remark 2.2.8. *Since any predictable process is optional, it follows from Theorem* 2.1.53 *[51], that for any $a \in \mathcal{A}^q$, the measure $\mu_a(A) = \langle 1_A, a\rangle$ is optional and then we have $\langle X, a\rangle = \langle \Pi^{op}(X), a\rangle$. That, together with* (2.2.11)*, yields*

$$\langle \Pi^{op}(X), a\rangle = \langle \Pi^{op}(X), \Pi^*(a)\rangle = \langle X, \Pi^*(a)\rangle, \qquad (X, a) \in \hat{\mathcal{R}}^p \times \hat{\mathcal{A}}^q. \quad (2.2.12)$$

Remark 2.2.9. *Let $Y \in L^p(\Omega, \mathcal{F})$ be a random variable. By Doob's Stopping Theorem it is easy to see that the optional projection of a constant random process $X_t = Y$, $\forall t \in [0, T]$ is the martingale $M_t := \mathbb{E}[Y|\mathcal{F}_t]$. Using* (2.2.12)*, it follows that for every $Y \in L^p$ and every $a = (a^l, a^r) \in \hat{\mathcal{A}}^q$ which is also adapted, one has*

$$\mathbb{E}\big[(a^l_T + a^r_T - a^r_0)Y\big] = \langle X, a\rangle = \langle M, a\rangle\,. \qquad (2.2.13)$$

Definition 2.2.6. *To every convex risk measure ρ on $\mathcal{R}^p$, one associates a convex risk measure on L^p, called the static risk,*

$$\bar{\rho}(Y) := \rho\bigg(\mathbb{E}\big[Y\big|\mathcal{F}_t\big]_{0\le t\le T}\bigg), \quad Y \in L^p,$$

and a static minimal penalty,

$$\overline{\rho^{\#}}(f) := \inf_{\{a\in\mathcal{D}_\sigma \,|\, \mathrm{Var}(a)=f\}} \rho^{\#}(a), \quad f \in D_\sigma.$$

Now we have the following theorem, which will be proven in the Appendix.

Theorem 2.2.3. *For every risk measure $\rho : \mathcal{R}^p \to \mathbb{R}$ the static minimal penalty equals the minimal static penalty i.e.*

$$\overline{\rho^{\#}} = (\bar{\rho})^{\#}.$$

Remark 2.2.10. *By Corollary 2.2.1, every coherent risk measure ρ on $\mathcal{R}^\infty$ with the Fatou property, can be identified with a subset $\mathcal{Q}$ of $\mathcal{D}_\sigma$. Let $\mathcal{P} = \mathrm{Var}(\mathcal{Q}) := \{\mathrm{Var}(a) : a \in \mathcal{Q}\}$. By relation* (2.2.13) *it is easy to see that for all $Y \in L^\infty$,*

$$\bar{\rho}(Y) = \sup_{f\in\mathcal{P}} \mathbb{E}\big[-fY\big]. \qquad (2.2.14)$$

2.3. Theoretical Results

We will now state our theoretical results. Their proofs are given in the Appendix.

In [50] it is shown that having the *Lebesgue* property for a convex risk measure with the *Fatou* property is equivalent to the weak compactness of lower contour sets of the conjugate function. In their proof, the author of [50] use the fact that for any uniformly integrable set $\mathcal{P} \subseteq L^1$ and uniformly bounded sequence Y_n converging in probability to Y we have:

$$\sup_{f\in\mathcal{P}} \mathbb{E}\big[Y_n f\big] \to \sup_{f\in\mathcal{P}} \mathbb{E}\big[Y f\big]. \tag{2.3.1}$$

In order to extend the *Lebesgue* property to bounded càdlàg process risk measures, we need to find an analog of (2.3.1) for the space of bounded càdlàg processes. Uniform integrability is relative compactness in the weak topology for L^1, so we could use the $\sigma(\mathcal{A}^1, \mathcal{R}^\infty)$ relatively compact set of $\mathcal{A}^1$ instead. Here it is worthwhile to mention that the Dunford-Pettis Theorem, which has the same spirit, states that for any two sequences f_n in L^1 and Y_n in the dual space L^∞, converging weakly to f, Y respectively, the sequence $\mathbb{E}[f_n Y_n]$ converges to $\mathbb{E}[fY]$. Therefore, knowing the $\sigma(\mathcal{A}^1, \mathcal{R}^\infty)$-compact subsets of $\mathcal{A}^1$ would allow us to characterize the *Lebesgue property for convex risk measures on* $\mathcal{R}^\infty$. *This characterization, when restricted to* L^∞, *yields the characterization in [50]. This can be carried out with the embedding* $i : L^\infty \to \mathcal{R}^\infty$ *defined as* $i(Y) = 1_{[T]}Y$. *On the other hand, we find that the compactness of a set* $\mathcal{Q}$ *in the topology* $\sigma(\mathcal{A}^1, \mathcal{R}^\infty)$ *is related to the compactness of* $\mathrm{Var}(\mathcal{Q})$, *the variation of* $\mathcal{Q}$.

In this section we start by characterizing compact subsets of $\mathcal{A}^q$ *with respect to the compact subsets of* L^q. *The first result is used to characterize compact sets of* $\mathcal{A}^q$. *This will be useful in applications.*

Theorem 2.3.1. *Let* $1 \le p, q \le \infty$ *be such that* $\frac{1}{p} + \frac{1}{q} = 1$, *and suppose that* $\mathcal{Q} \subset \mathcal{A}^q$. *The following conditions are equivalent:*

(C1) $\mathcal{Q}$ *is relatively compact in* $\sigma(\mathcal{A}^q, \mathcal{R}^p)$.

(C2) $\mathrm{Var}(\mathcal{Q})$ *is relatively compact in* $\sigma(L^q, L^p)$.

Furthermore, when $p = \infty$, *(C1) and (C2) are equivalent to*

(C3) $\mathcal{Q}$ *is bounded and for all* $\varepsilon > 0$ *there exists* $\eta > 0$ *such that for all* $X \in \mathcal{R}^\infty$ *bounded by 1 and with* $\mathbb{E}[X^*] \leq \eta$, *we have*

$$\sup_{a\in\mathcal{Q}} \langle |X|, a\rangle < \varepsilon. \tag{2.3.2}$$

The following corollary is an immediate consequence of Theorem 2.3.1.

Corollary 2.3.1. $\mathcal{Q} \subseteq \mathcal{A}^1$ *is* $\sigma(\mathcal{A}^1, \mathcal{R}^\infty)-$ *relatively compact if and only if* $\mathrm{Var}(\mathcal{Q})$ *is uniformly integrable.*

In the following discussions we consider that the risk measures always have a robust representation such as (2.2.2) *with* $\mathcal{D}_\sigma \cap \mathcal{A}^q$ *instead of* $\mathcal{D}_\sigma$. *By Theorem 2.2.1 for the case* $p = \infty$, *it is equivalent to assume that the convex risk measures have the Fatou property.*

Proposition 2.3.1. *For* $1 \leq p < \infty$, *every convex risk measure* $\rho : \mathcal{R}^p \to \mathbb{R}$ *having representation* (2.2.2) *also has the Lebesgue property.*

When $p = \infty$, *we have the following result.*

Theorem 2.3.2. *Let* $\rho : \mathcal{R}^\infty \to \mathbb{R}$ *be a convex risk measure with Fatou property. Then the following are equivalent:*

(L1) ρ *has Lebesgue property.*

(L2) For all $c \in \mathbb{R}$, $\{a \in \mathcal{A}^1;\ \rho^*(a) \leq c\}$ *is relatively compact in* $\sigma(\mathcal{A}^1, \mathcal{R}^\infty)$.

(L3) For all $c \in \mathbb{R}$, $\{a \in \mathcal{D}_\sigma;\ \rho^\#(a) \leq c\}$ *is relatively compact in* $\sigma(\mathcal{A}^1, \mathcal{R}^\infty)$.

(L4) ρ *always attains its maximum in* (2.2.3) .

(L5) $\bar{\rho}$ *has Lebesgue property.*

(L6) For all $c \in \mathbb{R}$, $\{f \in L^1;\ (\bar{\rho})^*(f) \le c\}$ *is relatively compact in* $\sigma(L^1, L^\infty)$.

(L7) For all $c \in \mathbb{R}$, $\{f \in D_\sigma;\ (\bar{\rho})^\#(f) \le c\}$ *is relatively compact in* $\sigma(L^1, L^\infty)$.

(L8) $\bar{\rho}$ *always attains its maximum in* (2.2.6).

We complete the section by stating a result which is a form of James' Theorem for the duality $(\mathcal{A}^q, \mathcal{R}^p)$. *Indeed, for the case* $1 < p < \infty$ *one can immediately see, by Remark 2.2.5, that the Theorem 2.3.3 is a form of James' Theorem. As for the case* $p = 1$, *it can be easily seen that Theorem 2.3.3 holds by direct application of the James' Theorem on duality* $(\hat{\mathcal{R}}^\infty, \hat{\mathcal{A}}^1)$ *and using the continuity of* Π^* *(Remark 2.2.7). What is not immediate and needs some justifications is the case* $p = \infty$, *which will be proven in the Appendix.*

Theorem 2.3.3 (James' Theorem for $(\mathcal{A}^q, \mathcal{R}^p)$)**.** *Let* $\mathcal{Q} \subseteq \mathcal{A}^q_+$ *be a convex,* $\sigma(\mathcal{A}^q, \mathcal{R}^p)$*-closed subset of* $\mathcal{A}^q$. *The set* $\mathcal{Q}$ *is compact in* $\sigma(\mathcal{A}^q, \mathcal{R}^p)$ *if and only if for each member* $X \in \mathcal{R}^p$ *it attains its supremum on* $\mathcal{Q}$.

2.4. Examples of Risk Measures with *Lebesgue* Property

In this section we present the first series of examples. In the sequel we will see how our results can help to figure out whether a convex risk measure has the Lebesgue property or not.

Before giving examples, we show how the Lebesgue property can be used in order to approximate the risk. Let X *be a random process in* $\mathcal{R}^\infty$. *A natural way of approximating this process within a time discretization is to construct the following sequence*

$$X_n = \sum_{i=0}^{2^n-1} 1_{[\frac{i}{2^n}T, \frac{i+1}{2^n}T)} X_{\frac{i}{2^n}T} + X_T. \tag{2.4.1}$$

It is clear that since X *is a càdlàg process,* $(X_n - X)^*$ *converges to zero in probability. Now, for a convex risk measure* ρ *with Lebesgue property we have that* $\rho(X_n) \to \rho(X)$. *This is no longer true if we only know that the risk measure* *has the Fatou property. Actually having the Fatou property, we can only say that $\rho(X_n)$ converges to $\rho(X)$ if X_n decreases to X. On the other hand, as a decreasing sequence, we cannot choose*

$$X'_n = \sum_{i=0}^{2^n-1} 1_{[\frac{i}{2^n}T, \frac{i+1}{2^n}T)} \sup_{[\frac{i}{2^n}T, \frac{i+1}{2^n}T)} X_t + X_T, \tag{2.4.2}$$

which is no longer adapted. As one can see, the Lebesgue property is a very strong assumption in approximating risk. In what follows, interestingly we will see that many important examples have the Lebesgue property which enables us to approximate.

In the following discussions, the first two examples are taken from [22]. The third one is the Snell envelope of a random process which is used in pricing an American option. In the fourth example we introduce for the first time a Cumulative-Stopping risk measure. This risk measure, besides having a very natural structure, provides us an exact formula for allocation when we deal with an α-stable random process.

In the sequel $\mathcal{P}_\sigma$ is a subset of $D_\sigma \cap L^q$ for $1 \leq q \leq \infty$. The coherent risk measure ρ_σ is defined on L^p as follows

$$\rho_\sigma(X) := \sup_{f \in \mathcal{P}_\sigma} \mathbb{E}\big[-fX\big]. \tag{2.4.3}$$

Example 2.4.1. *Let Θ be a set of stopping times and ρ be defined as follows*

$$\rho(Y) = \sup_{a \in \mathcal{Q}_\Theta} \langle -Y, a \rangle, \tag{2.4.4}$$

where $\mathcal{Q}_\Theta = \left\{ \big(0, \mathbb{E}[f|\mathcal{F}_\theta] 1_{t \geq \theta}\big) \middle| f \in \mathcal{P}_\sigma, \theta \in \Theta \right\}$.

For example, Θ can be a ruin time or the time that insurance surplus hits a specific barrier (see for instance [6]). Also, Θ can be the set of exercising times of an American option.

It is easy to see that

$$\rho(X) = \sup_{\theta \in \Theta} \rho_\sigma(X_\theta) \ , \ \forall X \in \mathcal{R}^p.$$

By (2.2.12) *and Remark 2.2.10, the static risk is calculated as*

$$\begin{aligned}\bar{\rho}(Y) &= \sup_{a\in\mathcal{Q}}\langle -(\mathbb{E}[Y|\mathcal{F}_t])_{t\in[0,T]}, a\rangle \\ &= \sup_{f\in\mathcal{P}_\sigma,\theta\in\Theta} \mathbb{E}\Big[\mathbb{E}[-Y|\mathcal{F}_\theta]f\Big] \\ &= \sup_{f\in\mathcal{P}_\sigma,\theta\in\Theta} \mathbb{E}\Big[-Y\mathbb{E}[f|\mathcal{F}_\theta]\Big].\end{aligned}$$

According to Theorem 2.2.2, when $p=\infty$, $\bar{\rho}$ has Lebesgue property iff

$$\Big\{\mathbb{E}[f|\mathcal{F}_\theta]\Big|\theta\in\Theta\,,\ f\in\mathcal{P}_\sigma\Big\}$$

is uniformly integrable. Therefore, by Theorem 2.3.2, ρ has Lebesgue property iff the above set is uniformly integrable. In particular, it has Lebesgue property when $\mathcal{P}_\sigma$ is uniformly integrable. In other words, ρ has Lebesgue property if ρ_σ does.

Example 2.4.2. *For any random variable $f\in\mathcal{P}_\sigma\subseteq D_\sigma\cap L^q$ (for some $1\le q\le\infty$) and stopping time $\theta\in\Theta$, define the random process f_θ as follows:*

$$f_\theta(t) = \begin{cases} \frac{t}{\theta}\mathbb{E}[f|\mathcal{F}_t] & t\le\theta, \\ \\ \mathbb{E}[f|\mathcal{F}_\theta] & \textit{otherwise}. \end{cases} \tag{2.4.5}$$

Then, on $\mathcal{R}^p$, let

$$\rho(X) = \sup_{a\in\mathcal{Q}}\langle -X, a\rangle,$$

where $\mathcal{Q} = \Big\{(f_\theta,0)\Big|f\in\mathcal{P}_\sigma\,,\ \theta\in\Theta\Big\}$. It is easy to see that

$$\begin{aligned}\rho(X) &= \sup_{\theta\in\Theta}\rho_\sigma\bigg(\frac{1}{\theta}\int_0^\theta X_t dt\bigg), \\ \mathrm{Var}(\mathcal{Q}) &= \Big\{\mathbb{E}[f|\mathcal{F}_\theta]\ \Big|\ f\in\mathcal{P}_\sigma\,,\ \theta\in\Theta\Big\}, \\ \bar{\rho}(Y) &= \sup_{f\in\mathcal{P}_\sigma\,,\theta\in\Theta}\mathbb{E}\Big[-Y\mathbb{E}[f|\mathcal{F}_\theta]\Big]\ ,\ \textit{for } Y\in L^p.\end{aligned}$$

By part (C2) of Theorem 2.3.1 and (*L*3) *of Theorem 2.3.2, ρ has Lebesgue property iff*

$$\mathrm{Var}(\mathcal{Q}) = \Big\{\mathbb{E}[f|\mathcal{F}_\theta]\ \Big|\ f\in\mathcal{P}_\sigma\,,\ \theta\in\Theta\Big\}$$

is uniformly integrable when $p = \infty$. Also it has Lebesgue property if $\mathcal{P}_\sigma$ is uniformly integrable

Example 2.4.3 (Snell Envelope and American Option Price Stability). *Let $X \in \mathcal{R}^\infty$ and $S \le T$ be a stopping time. Let*

$$\Theta_S = \big\{\theta \ge S \big| \theta \text{ is } [0,T]\text{-value stopping time}\big\}.$$

Set

$$\begin{aligned}\rho_S(X) &= \operatorname*{ess\,sup}_{a \in \mathcal{Q}_S} \langle -X, a\rangle \\ &= \operatorname{ess\,sup}\left\{\mathbb{E}\big[-X_\theta \big| \mathcal{F}_S\big]\middle| \theta \in \Theta_S\right\}.\end{aligned}$$

The process $\rho_t(X)$ is the smallest super-martingale larger than $-X$ which is called the Snell envelope of $-X$, see, e.g., [22].

Now for any measurable set $A \in \mathcal{F}_S$ define

$$\rho_S^A(X) = \mathbb{E}\big[\rho_S(X) 1_A\big]. \tag{2.4.6}$$

It is exactly equivalent to put $\mathcal{P}_\sigma = \left\{\frac{1}{\mathbb{P}(A)} 1_A\right\}$ and $\Theta = \Theta_S$ in Example 1. From Example 1 we know that ρ_S^A has the Lebesgue property. Since the choice of $A \in \mathcal{F}_S$ is arbitrary, then by (2.4.6), we have that for each stopping time S the Snell envelope $\rho_S(X)$ is continuous in the weak star topology. In particular, setting $\rho_t = \rho_t^\Omega$, then $\rho_t(X_n) \to \rho_t(X)$ when $(X_n - X)^* \to 0$. This shows how one can approximate the price of an American option in continuous time by time discretization.

Example 2.4.4 (Cumulative-Stopping Risk). *Let ρ_σ be a risk measure on L^p. A natural way to assess the risk of a random process is the average of the risk over the time interval, i.e. $\frac{1}{T}\int_0^T \rho_\sigma(X_s)ds$. On the other hand, let us suppose that there exists a stopping time (or a general random time) which shows some crucial moments, important for the risk user. Then a way to measure the risk of a random process X in $\mathcal{R}^p$ is to calculate*

$$\rho(X) = \int_0^T \rho_\sigma(X_s) f_\theta(s) ds, \tag{2.4.7}$$

where f_θ is the density function of θ. This new convex risk measure is called the Cumulative-Stopping risk.

In fact, for any measure μ *on* $[0, T]$,

$$\int_0^T \rho_\sigma(X_s)\mu(ds)$$

will work and it is a mixture risk measure.

It is not very difficult to see that when the risk measure ρ_σ *is* $\sigma(L^p, L^q)$*-lower semi-continuous then* ρ *is also lower semi-continuous. It means that when* ρ_σ *has a representation like* (2.2.6) *then* ρ *has a representation like* (2.2.2) *(with* $\mathcal{D}_\sigma \cap \mathcal{A}^q$ *instead of* $\mathcal{D}_\sigma$*). On the other hand, when* $p = \infty$*, the convex risk* ρ *has the Lebesgue property iff* ρ_σ *does. Actually this follows from part (L5) of Theorem 2.3.2.*

2.5. Applications to the Capital Allocation Problem

In this section, we give an application of Theorem 2.3.2 to allocation of risk capital. This problem for one-period coherent risk measures was discussed in [33], where the weak star sub-gradient of a coherent risk measure was defined. It was shown that the existence of a solution for the capital allocation problem is equivalent to having a nonempty sub-gradient. James' Theorem played a key role in showing that the weak sub-gradient is not empty. In our setting, Theorem 2.3.3 plays almost the same role. The same allocation problem for dynamic coherent risk measures for discrete times was studied in [24]. For coherent allocation of risk capital, see [35] and the references therein.

We begin by recalling the definition of capital allocation. For more details see [32], [11] and [17].

In general, let $X_1, ..., X_N$ *be* N *random processes in* $\mathcal{R}^p$ *representing* N *financial positions, for example, the values of* N *departments of a firm. The total capital required to face the risk of* $X_1 + \cdots + X_N$ *is* $\rho(\sum_{i=1}^{N} X_i) = k$*. We want to find a "fair" allocation* $(k_1, \ldots, k_N)$ *so that* $k_1 + \cdots + k_N = k$.

Definition 2.5.1. *An allocation* $(k_1, \ldots, k_N)$ *with* $k = k_1 + \cdots + k_N$ *is called fair in fuzzy game approach (simply an allocation) if for all* α_j , $j = 1, \ldots, N$, $0 \leq \alpha_j \leq 1$

we have

$$\sum_j \alpha_j k_j \le \rho\left(\sum_j \alpha_j X_j\right).$$

Before moving on with our discussion we define the weak sub-gradient.

Definition 2.5.2. *For a function $\rho : \mathcal{R}^p \to \mathbb{R}$, the weak sub-gradient (in this article simply sub-gradient) of ρ at X is defined by*

$$\triangledown\rho(X) := \{a \in \mathcal{A}^q | \rho(X+W) \ge \rho(X) + \langle W, a\rangle, \quad \forall W \in \mathcal{R}^p\}. \tag{2.5.1}$$

When $p = \infty$ this set can be empty but for $p \neq \infty$ this set is always nonempty [59, Proposition 3.1].

We have the following extension of Theorem 17, Section 8.2 [32] without proof. Actually if one looks at the proof of Theorem 17, Section 8.2 [32], every part of the proof can be stated with X as a random process instead of random variable.

Theorem 2.5.1. *Let ρ be a coherent risk measure with representation* (2.2.4) *given by a family $\mathcal{Q} \subseteq \mathcal{D}_\sigma \cap \mathcal{A}^q$. Then $a \in \triangledown\rho(X)$ iff $-a \in \mathcal{Q}$ and $\rho(X) = \langle -X, -a\rangle = \langle X, a\rangle$.*

As a direct consequence of Theorems 2.3.2, 2.5.1 and 2.3.3, we have

Theorem 2.5.2. *Let $\rho : \mathcal{R}^p \to \mathbb{R}$ be a coherent risk measure with representation* (2.2.4) *when $\mathcal{Q} \subseteq \mathcal{D}_\sigma \cap \mathcal{A}^q$. The the following conditions are equivalent:*

- $\triangledown\rho(X) \neq \emptyset$, $\forall X \in \mathcal{R}^p$*;*
- $\mathcal{Q}$ *is* $\sigma(\mathcal{A}^q, \mathcal{R}^p)$*-compact;*
- $\mathrm{Var}(\mathcal{Q})$ *is* $\sigma(L^q, L^p)$*-compact;*
- ρ *has the Lebesgue property;*
- $\bar{\rho}$ *has the Lebesgue property.*

Finally, we can state the solution of the optimal allocation problem, using Theorems 2.5.1, 2.3.3 and 2.5.2.

Theorem 2.5.3. *If $X = X_1 + \cdots + X_N$ and if $-a \in \triangledown\rho$, then the allocation $k_i = \langle -X_i, a\rangle$ is a fair allocation.*

2.5.1. Calculating the Sub-gradient

Before giving the examples we calculate the sub-gradient of the risk measure constructed in Example 1 by considering $\Theta = \{\theta\}$. Again we consider a subset

$\mathcal{P}_\sigma \subseteq D_\sigma \cap L^q$ and we let $\mathcal{Q} = \big\{\big(0, \mathbb{E}[f|\mathcal{F}_\theta]\big)1_{\theta \geq t}\big| f \in \mathcal{P}_\sigma\big\}$. It is easy to see that:

$$\rho(X) = \sup_{a \in \mathcal{Q}} \langle -X, a\rangle = \sup_{f \in \mathcal{P}_\sigma} \mathbb{E}[-X_\theta f] = \rho_\sigma(X_\theta). \tag{2.5.2}$$

Now, consider that $-f \in \bigtriangledown \rho_\sigma(X_\theta)$. We have

$$\rho(X) = \rho_\sigma(X_\theta) = \mathbb{E}\big[-X_\theta f\big] = \big\langle -X, \big(0, \mathbb{E}\big[f|\mathcal{F}_\theta\big]\big)1_{t \geq \theta}\big\rangle. \tag{2.5.3}$$

Since $\big(0, \mathbb{E}\big[f|\mathcal{F}_\theta\big]\big)1_{t\geq\theta} \in \mathcal{Q}$, then by Theorem 2.5.1, $\big(0, \mathbb{E}\big[-f|\mathcal{F}_\theta\big]\big)1_{t\geq\theta} \in \bigtriangledown\rho(X)$. Hence

$$\left\{\big(0, \mathbb{E}[-f|\mathcal{F}_\theta]\big)1_{t\geq\theta}\,\middle|\, -f \in \bigtriangledown \rho_\sigma(X_\theta)\right\} \subseteq \bigtriangledown\rho(X). \tag{2.5.4}$$

On the other hand, if we take $-a \in \bigtriangledown\rho(X)$ then from Theorem 2.5.1 it turns out that $a = \big(0, \mathbb{E}[f|\mathcal{F}_\theta]1_{t\geq\theta}\big)$. Therefore,

$$\rho_\sigma(X_\theta) = \rho(X) = \big\langle -X, \big(0, \mathbb{E}[f|\mathcal{F}_\theta]1_{t\geq\theta}\big)\big\rangle = \mathbb{E}[-X_\theta f]. \tag{2.5.5}$$

Since $f \in \mathcal{P}_\sigma$, this shows that $-f \in \bigtriangledown\rho_\sigma(X_\theta)$, which in turn yields

$$\bigtriangledown\rho(X) \subseteq \big\{\big(0, \mathbb{E}[-f|\mathcal{F}_\theta]\big)1_{t\geq\theta}\big| -f \in \bigtriangledown\rho_\sigma(X_\theta)\big\}.$$

Combining that with (2.5.4), we end up with

$$\bigtriangledown\rho(X) = \left\{\big(0, \mathbb{E}[-f|\mathcal{F}_\theta]\big)1_{t\geq\theta}\,\middle|\, -f \in \bigtriangledown\rho_\sigma(X_\theta)\right\}. \tag{2.5.6}$$

2.5.2. Examples of Capital Allocation

Here we present some examples which could be used in real life problems. They are mostly designed to consider the problem of capital allocation for departments of a firm such as an insurance company. The risk processes which make the core of the insurance risk theory is one of the main subjects we frequently look at. In particular, we solve the whole problem for capital allocation for α-stable risk processes. In this section we present the following four examples. Example 5 is a general problem of Quantile Based Allocation, which gives the allocation in term of the process at maturity and the stopping time. This example is a basis for the two next examples. In Example 6 we consider the same problem of quantile based allocation for a Brownian motion case. In Example 7 we pose an insurance risk problem from the point of view of an insurance company. In Example 8 we study *the same problem as in Example 7, with Cumulative-Stopping risk measure and also an α-stable process.*

Example 2.5.1 (Quantile Based Allocation)**.** *Let $X_1, \dots, X_N$ be random processes representing the evolution of the future values of N departments. Let $X = X_1 + \dots + X_N$, $\Theta = \{\theta\}$, and*

$$\mathcal{P}_\sigma = \left\{ h \in L^1(\Omega, \mathcal{F}_T, \mathbb{P})_+ \,\middle|\, \mathbb{E}[h] = 1, 0 \le h \le \frac{1}{\alpha} \right\},$$

for some confidence level $0 < \alpha < 1$. Here ρ_σ is $AVaR_\alpha^{\mathcal{F}_T}$. Since $\mathcal{P}_\sigma \subseteq L^\infty$ then ρ_σ is a risk measure on L^1 and the corresponding measure ρ is defined for $\mathcal{R}^1$. From [32, Section 8], we know that if X_θ is continuous then

$$\bigtriangledown AVaR_\alpha(X_\theta) = \left\{ -\frac{1}{\alpha} 1_A \right\}, \tag{2.5.7}$$

where $A = \big\{ X_\theta \le q_\alpha(X_\theta) \big\}$.

From Theorem 2.5.3 and Example 4, the allocation $(k_1, \dots, k_n)$ is given by:

$$k_i = -\frac{1}{\alpha} \mathbb{E}\Big[X_{i,\theta} 1_A \Big], \tag{2.5.8}$$

Now let $\mathbb{Q}$ be an equivalent measure to $\mathbb{P}$ under which X is a martingale. Then we have

$$A = \left\{ \frac{\mathbb{E}[X_T \frac{d\mathbb{Q}}{d\mathbb{P}} | \mathcal{F}_\theta]}{\mathbb{E}[\frac{d\mathbb{Q}}{d\mathbb{P}} | \mathcal{F}_\theta]} \le q_\alpha \left(\frac{\mathbb{E}[X_T \frac{d\mathbb{Q}}{d\mathbb{P}} | \mathcal{F}_\theta]}{\mathbb{E}[\frac{d\mathbb{Q}}{d\mathbb{P}} | \mathcal{F}_\theta]} \right) \right\}. \tag{2.5.9}$$

which gives A in terms of X at maturity and stopping time θ.

Example 2.5.2. *Let $W = (W^1, \dots, W^M)$ be an M dimensional of independent Brownian motions. In the Example 2.5.1 let*

$$d\vec{X}_t = \mu_t dt + \sigma dW, \tag{2.5.10}$$

where each (possibly random) component μ_t^i of μ_t is a positive function satisfying Novikov's conditions and σ is a deterministic $N \times M$ matrix. By applying Doob's inequality for martingales, one can see that $X_i \in \mathcal{R}^1$. Actually since the function $x \mapsto |x|$ is a convex function then $|W_t|$ is a sub-martingale. Then by Doob's martingale inequality we have

$$\mathbb{P}\left[W^* = \sup_{0 \le t \le T} |W_t| \ge c \right] \le \frac{\mathbb{E}\left[|W_T|^2 \right]}{c^2}.$$

Now

$$\mathbb{E}\left[W^*\right] = \int_0^\infty \mathbb{P}\left[W^* \geq c\right] dc \leq 1 + \mathbb{E}\left[|W_T|^2\right] \int_1^\infty \frac{1}{c^2} dc < \infty.$$

Note that

$$dX = \left(\sum_{i=1}^{N} \mu_t^i \right) dt + \sum_{i=1}^{N} \sum_{j=1}^{M} \sigma_{ij} dW^j, \tag{2.5.11}$$

which can be rewritten as

$$dX = \tilde{\mu}_t dt + \tilde{\sigma} d\tilde{W}, \tag{2.5.12}$$

where $\tilde{\mu} = \sum_{i=1}^{n} \mu_t^i$, $\tilde{\sigma} = \left(\sum_{j=1}^{M} \left(\sum_{i=1}^{N} \sigma_j^i \right)^2 \right)^{1/2}$ *and* $\tilde{W}$ *is a Brownian motion.*

So we have $\frac{X_t}{\tilde{\sigma}} = \frac{1}{\tilde{\sigma}} \int_0^t \tilde{\mu}_s ds + \tilde{W}_t$*. By Girsanov's Theorem* $\frac{X_t}{\tilde{\sigma}}$ *is a martingale under the measure* $\mathbb{Q}$ *defined as*

$$\mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \middle| \mathcal{F}_t \right] = \exp\left(- \int_0^t \frac{\tilde{\mu}_s}{\tilde{\sigma}} d\tilde{W} + \frac{1}{2} \int_0^t \left(\frac{\tilde{\mu}_s}{\tilde{\sigma}} \right)^2 ds \right).$$

Using (2.5.8) *we have*

$$k_i = -\frac{1}{\alpha} \mathbb{E}\left[1_A \; X_{i,\theta} \right], \tag{2.5.13}$$

where

$$A = \Bigg\{ \mathbb{E}\left[X_T \exp\left(- \int_\theta^T \frac{\tilde{\mu}_s}{\tilde{\sigma}} d\tilde{W} + \frac{1}{2} \int_\theta^T \left(\frac{\tilde{\mu}_s}{\tilde{\sigma}} \right)^2 ds \right) \middle| \mathcal{F}_\theta \right]$$
$$\leq q_\alpha \left(X_T \exp\left(- \int_\theta^T \frac{\tilde{\mu}_s}{\tilde{\sigma}} d\tilde{W} + \frac{1}{2} \int_\theta^T \left(\frac{\tilde{\mu}_s}{\tilde{\sigma}} \right)^2 ds \right) \right) \Bigg\},$$

which gives A *in terms of* X *at maturity and stopping time* θ.

Example 2.5.3. *Suppose that there is an insurance company consisting of* N *departments. The surplus of the i-th department is denoted with* $X_{i,t}$, $1 \leq i \leq N$. *Suppose that* $\vec{X} = (X_{1,t}, \ldots, X_{N,t})$ *is modeled by the following process:*

$$\vec{X}_t = \vec{c}(t) + \vec{L}_t, \tag{2.5.14}$$

where $\vec{c}(t)$ *is an increasing process and* $\vec{L}_t$ *is an* N*-dimensional process* $\vec{L}_t \in (\mathcal{R}^1)^N$. *This model appears in the insurance risk theory when in general* L_t *is a Lévy process with non-positive jumps. This is what one calls the generalized Cramèr-Lundberg process. In more detail,* $c_i(t)$ *represents the premium received by the i-th department while* $L_{i,t}$ *represents the claims (see, e.g., [56] and [52]). For instance, using Doob's martingale inequality, (similar to what we have done* *in the previous example) one sees that an α-stable Lévy process with parameters $1 < \alpha < 2$ and $\beta = -1$ is a Lévy process without positive jumps in $\mathcal{R}^1$. The characteristic function of all Lévy processes without positive jumps is given as*

$$\mathbb{E}\big[e^{i\vec{\lambda}\cdot\vec{X}_1}\big] \;=\; \exp\left\{i\vec{a}\cdot\vec{\lambda} + \frac{1}{2}\vec{\lambda}^T Q\vec{\lambda} + \int_{(-\infty,0)^N}\left(e^{i\vec{\lambda}\cdot\vec{x}} - 1 - i\vec{\lambda}\cdot\vec{x}1_{\{|\vec{x}|<1\}}\right)\Pi(d\vec{x})\right\},$$

where $i^2 = -1$, Q is a positively definite $N\times N$ matrix, $\vec{a}$ is an N-dimensional drift vector and Π is a measure on $(-\infty,\infty)^N$ for which $\int_{\mathbb{R}^N}(1\wedge|x|^2)\Pi(d\vec{x}) < \infty$. From this last relation one can see that the process $X = X_1 + \cdots + X_N$ equals $c(t) + L_t$, where $c(t) = c_1(t) + \cdots + c_N(t)$ is an increasing premium and $L_t = \sum_j L_{j,t}$ is a claim process without positive jumps. Let $\mu_j = \mathbb{E}[L_{j,1}]$. It is clear that $L_{j,t} - \mu_j t$ is a martingale. Let $d_j(t) = c_j(t) + \mu_j t$ and $d(t) = \sum d_j(t)$. Then $X_t - d(t)$ and $X_{j,t} - d_j(t)$ for $1 \leq j \leq N$ are martingales.

Now the quantile allocation is given by (2.5.8) as follows

$$k_j = -\frac{1}{\alpha}\mathbb{E}\Big[1_A\big(X_{j,T} - d(T) + d(\theta)\big)\Big], \tag{2.5.15}$$

where

$$A = \left\{\mathbb{E}[X_T|\mathcal{F}_\theta] + d_j(\theta) \leq q_\alpha\left(\mathbb{E}[X_T|\mathcal{F}_\theta] + d_j(\theta)\right)\right\}. \tag{2.5.16}$$

An interesting observation about (2.5.15) is the incorporation of the stopping time in the allocation. It shows how the allocation is jointly affected by X_T and θ. For example we see what happens when $\mathcal{F}_\theta$ is independent from X_T: in such a case we have

$$k_j = -d_j(T) + d(T) - \frac{1}{\alpha}\mathbb{E}[\theta|\theta \leq q_\alpha(\theta)],$$

where we can see how everything depends only on the stopping time.

Example 2.5.4 (Cumulative-Stopping Allocation)**.** *In this example again we consider an insurance company with N departments. We set up the model of the previous example when X is a multivariate α-stable process and the risk measure is a Cumulative-Stopping risk. For that, let $(Z_{1,t},\ldots,Z_{N,t})$ be a N-dimensional α-stable Lévy processes with $1 < \alpha < 2$. By Doob's martingale inequality we* *know that* $Z_i \in \mathcal{R}^1$. *For some positive numbers* c^i *and positive numbers,* a_1^i, a_2^i, $i = 1, \ldots, N$ *let*

$$X_{i,t} = c^i t + a_1^i Z_{1,t} + \cdots + a_N^i Z_{N,t}\ ,\ i = 1, \ldots, N.$$

In this example we suppose that the company is concerned with some financial position made in the market. There are some crucial moments in which this financial position is at risk. These moments are modeled with a random time θ. *The company uses the risk measure* AVaR_a *to asses the risk at each single time* $t \in [0, T]$. *(To avoid any confusion between the* α*'s in the definition of risk* AVaR_α *and the* α *in* α*-stable process we use the notation* AVaR_a *for some* $0 < a < 1$ *instead of* AVaR_α*.) We would like to find the risk allocated to each department with respect to the Cumulative-Stopping risk* $\mathrm{AVaR}_a^{\theta,CS}(X) = \int_0^T \mathrm{AVaR}_a(X_s) f_\theta(s) ds$.

Let $(k'_{1,t}, \ldots, k'_{N,t})$ *be an allocation for the static problem* $X_t = X_{1,t} + \cdots + X_{N,t}$ *using the risk measure* AVaR_a. *We define the random variables* $K_{i,\theta}(\omega) = k'_{i,\theta(\omega)}$ *and then we define* $k_{i,\theta} = \mathbb{E}\left[K_{i,\theta}\right]$ *for* $i = 1, \ldots, N$. *For* $0 \leq \alpha_1, \ldots, \alpha_N \leq 1$ *we have:*

$$\begin{aligned}
\alpha_1 k_{1,\theta} + \ldots, \alpha_N k_{N,\theta} &= \mathbb{E}\left[\alpha_1 K_{1,\theta} + \cdots + \alpha_N K_{N,\theta}\right] \\
&= \int_0^T (\alpha_1 k'_{1,s} + \cdots + \alpha_N k'_{N,s}) f_\theta(s) ds \\
&\leq \int_0^T \mathrm{AVaR}_a(\alpha_1 X_{1,s} + \cdots + \alpha_N X_{N,s}) f_\theta(s) ds \\
&= \mathrm{AVaR}_a^{\theta,CS}(\alpha_1 X_1 + \cdots + \alpha_N X_N),
\end{aligned}$$

and the inequality is equality when $\alpha_1 = \cdots = \alpha_N = 1$. *According to definition* $(k_{1,\theta}, \ldots, k_{N,\theta})$ *is an allocation for* $(X_1, \ldots, X_N)$.
Let $(l'_{1,t}, \ldots, l'_{N,t})$ *be an allocation for* $(Z_{1,t}, \ldots, Z_{N,t})$. *It is clear that*

$$(l'_{1,t}, \ldots, l'_{N,t}) = (c_1 t + k'_{1,t}, \ldots, c_N t + k'_{N,t}).$$

Since $Z_t = Z_{1,t} + \cdots + Z_{N,t}$ *has the scaling property (i.e.* $Z_t \overset{d}{=} t^{\frac{1}{\alpha}} Z_1$*) and* AVaR_a *is positively homogeneous and law invariant risk measure, we have that* $l'_{i,t} = t^{\frac{1}{\alpha}} l'_{i,1}$. *Therefore we have*

$$k'_{i,t} = -c_i t + t^{\frac{1}{\alpha}} (k'_{i,1} + c_i) \ \textit{ for }\ i = 1, \ldots, N.$$

Now

$$\begin{aligned} k_{i,\theta} &= \mathbb{E}[K_{i,\theta}] \\ &= \int_0^T k'_{i,s} f_\theta(s) ds \\ &= \int_0^T \left(-c_i t + t^{\frac{1}{\alpha}} (k'_{i,1} + c_i)\right) f_\theta(s) ds \\ &= -c_i \mathbb{E}[\theta] + (k'_{i,1} + c_i) \mathbb{E}\left[\theta^{\frac{1}{\alpha}}\right]. \end{aligned}$$

We also know that $k'_{i,1} = -\mathbb{E}\big[X_{i,1} | X_{1,1} + \cdots + X_{N,1} \leq q_a(X_{1,1} + \cdots + X_{N,1})\big]$, *which yields:*

$$k_{i,\theta} = -c_i \mathbb{E}[\theta] - \left(\mathbb{E}\left[X_{i,1} \middle| X_{1,1} + \cdots + X_{N,1} \leq q_a(X_{1,1} + \cdots + X_{N,1})\right] + c_i\right) \mathbb{E}\left[\theta^{\frac{1}{\alpha}}\right].$$

Here one can see that using an α-stable model in Example 8, along with Cumulative-Stopping risk measure, yields an allocation $k_{i,\theta}$ that is proportional to the allocation $k_{i,1}$. The constant of proportionality is nothing but the expectation of the stopping time to the power $1/\alpha$. This shows that the later the event associated to θ occurs, the larger the effect on $k_{i,1}$ is. Moreover, this shows an inverse relation between the parameter α and the effect of the second term.

2.6. Proofs of the Theoretical Results

2.6.1. Proof of Theorem 2.3.1

We split the proof into two main parts according as $p \neq \infty$ and $p = \infty$.

Proof for $p \neq \infty$.

$(C2) \Leftrightarrow (C1)$

Since by Remark 2.2.5 $\hat{\mathcal{A}}^q$ is the topological dual of $(\hat{\mathcal{R}}^p, \sigma(\hat{\mathcal{R}}^p, \hat{\mathcal{A}}^q))$, $\hat{\mathcal{A}}^q$ is endowed with the weak topology. Therefore, $\mathcal{Q}$ is relatively compact iff it is bounded and the latter is true iff $\mathrm{Var}(\mathcal{Q})$ is bounded. In other words, $\mathcal{Q}$ is relatively compact in $\sigma(\hat{\mathcal{A}}^q, \hat{\mathcal{R}}^p)$ iff $\mathrm{Var}(\mathcal{Q})$ is relatively compact in $\sigma(L^q, L^p)$. Now the assertion $(C2) \Leftrightarrow (C1)$ is true because of the continuity of $\Pi^* : \hat{\mathcal{A}}^q \to \mathcal{A}^q$ (Remark 2.2.7).*

Proof for $p = \infty$.

$(C2) \Rightarrow (C1)$ *We define a topology on $\mathcal{R}^\infty$, generated by a family of semi-norms. For any weakly relatively compact subset $\mathcal{P}$ in L^1 let*

$$V(\mathcal{P}) := \left\{ a \in \mathcal{A}^1 \middle| \exists f \in \mathcal{P} \ s.t. \mathrm{Var}(a) \leq |f| \right\}$$

and define the associated semi-norm for $\mathcal{P}$ on $\mathcal{R}^\infty$ with

$$P_{\mathcal{P}}(X) = \sup_{a \in V(\mathcal{P})} \langle X, a \rangle.$$

This topology is compatible with the vector structure because all $V(\mathcal{P})$'s are bounded. We denote this topology by σ^1. Let $(\mathcal{R}^\infty)'$ be the dual of $\mathcal{R}^\infty$ with respect to the topology σ^1. It is clear that $\mathcal{A}^1 \subseteq (\mathcal{R}^\infty)'$. We want to show that $\mathcal{A}^1 = (\mathcal{R}^\infty)'$. Let μ be an arbitrary element of $(\mathcal{R}^\infty)'$ and X_n be a non-negative sequence of uniformly bounded members in $\mathcal{R}^\infty$ such that $(X_n)^* \xrightarrow{\mathbb{P}} 0$. By (2.3.1), we have

$$0 \leq P_{\mathcal{P}}(X_n) \leq \sup_{f \in \mathcal{P}} \mathbb{E}[(X_n)^* |f|] \to 0. \tag{2.6.1}$$

(2.6.1) *implies that $X_n \xrightarrow{\sigma^1} 0$ and therefore $\mu(X_n) \to 0$. This shows that μ is finitely additive. Also from* (2.6.1) *it yields that the functional μ is order bounded (i.e., for every W, $\sup_{\{U \leq W\}} \mu(U) < \infty$). Since $\mathcal{R}^\infty$ is a Riesz space, from the general theory of Riesz spaces μ can be decomposed into the difference of its positive and negative parts (for example see [3], Theorem 3.3). Let μ^+ be* *the positive part. By definition of the positive part, for any* $X \geq 0$, $\mu^+(X) = \sup_{0\leq W\leq X} \mu(W)$*. Let* X_n *be a positive and decreasing sequence for which* $(X_n)^* \downarrow 0$ *in probability. Let* $0 \leq W_n \leq X_n$ *be such that* $\mu^+(X_n) \leq \mu(W_n) + \frac{1}{n}$*. Since* $(W_n)^* \xrightarrow{\mathbb{P}} 0$*, by (2.6.1) we get that*

$$0 \leq \mu^+(X_n) \leq \mu(W_n) + \frac{1}{n} \to 0.$$

Given Theorem 2 of chapter VII in [34], one deduces that $\mu^+ \in \mathcal{A}^1$*. Similarly* $\mu^- \in \mathcal{A}^1$ *and therefore* $\mu \in \mathcal{A}^1$*. This completes the proof that* $\mathcal{A}^1 = (\mathcal{R}^\infty)'$*.*
The Corollary to Mackey's Theorem 9, section 13, chapter 2 [44] leads us to $\sigma^1 \subseteq \tau(\mathcal{R}^\infty, \mathcal{A}^1)$*, where* $\tau(\mathcal{R}^\infty, \mathcal{A}^1)$ *is the Mackey topology. Just to recall, Mackey's topology is a topology generated with a basis of open sets around the origin defined as*

$$\left\{X \in \mathcal{R}^\infty \middle| \sup_C \langle X, a\rangle < 1\right\},$$

for all $\sigma(\mathcal{A}^1, \mathcal{R}^\infty)$*-compact subsets* $C \subseteq \mathcal{A}^1$*.*

Let $\mathcal{P}$ *be a* $\sigma(L^1, L^\infty)$ *relatively compact subset of* L^1*. By definition, the set* $\{X | P_{\mathcal{P}}(X) < 1\}$ *is an open set in* σ^1*. Since* $\sigma^1 \subseteq \tau$*, this set is also an open set in* τ*. Therefore, there exists a* $\sigma(\mathcal{A}^1, \mathcal{R}^\infty)$*-compact set* C *such that* $\{X | \sup_C \langle X, a\rangle < 1\} \subseteq \{X | P_{\mathcal{P}}(X) < 1\}$*. By polarity (which is decreasing with respect to inclusion) we have that* $\{X | P_{\mathcal{P}}(X) \leq 1\}^\circ \subseteq \{X | P_{\mathcal{P}}(X) < 1\}^\circ \subseteq \{X | \sup_{a\in C} \langle X, a\rangle < 1\}^\circ$*. From the generalized Bourbaki-Alaoglu Theorem we know that the polar set of every open set in* $(\mathcal{R}^\infty, \sigma^1)$*, which we know has* $\mathcal{A}^1$ *as its dual, is* $\sigma(\mathcal{A}^1, \mathcal{R}^\infty)$*-compact. Therefore,* $\{X | \sup_C \langle X, a\rangle < 1\}^\circ$ *is* $\sigma(\mathcal{A}^1, \mathcal{R}^\infty)$*-compact. Since* $\{X | P_{\mathcal{P}}(X) \leq 1\}^\circ \subseteq \{X | \sup_C \langle X, a\rangle < 1\}^\circ$ *then* $\{X | P_{\mathcal{P}}(X) \leq 1\}^\circ$ *is* $\sigma(\mathcal{A}^1, \mathcal{R}^\infty)$*-relatively compact. By definition of polarity it is clear that* $V(\mathcal{P}) \subseteq \{X | P_{\mathcal{P}}(X) \leq 1\}^\circ$*, which yields that* $V(\mathcal{P})$ *is* $\sigma(\mathcal{A}^1, \mathcal{R}^\infty)$*-relatively compact. Now let* $\mathcal{P} = \mathrm{Var}(\mathcal{Q})$*, since* $\mathcal{Q} \subseteq V(\mathrm{Var}(\mathcal{Q}))$*, one concludes that* $\mathcal{Q}$ *is* $\sigma(\mathcal{A}^1, \mathcal{R}^\infty)$*-relatively compact.*

$(C1) \Rightarrow (C3)$. *Let*

$$G = \left\{ X \in \mathcal{R}^\infty \middle| \mathbb{E}[X^*] \leq 1\,,\ X \text{ is bounded by 1} \right\}.$$

Thus, G is a bounded set in the topology $\sigma(\mathcal{R}^\infty, \mathcal{A}^1)$. Indeed, this is true since for every $a \in \mathcal{A}^1$ we have $|\langle X, a\rangle| \leq \mathbb{E}[X^*]\mathrm{Var}(a) \leq \mathrm{Var}(a)$. This implies that for every relatively compact subset $\mathcal{Q}$ in $\sigma(\mathcal{A}^1, \mathcal{R}^\infty)$,

$$\sup_{X \in G} (\sup_{a \in \mathcal{Q}} |\langle X, a\rangle|) =: L < \infty. \tag{2.6.2}$$

Indeed, $X \mapsto \sup_{a \in \mathcal{Q}} \langle X, a\rangle$ is a semi-norm from which the Mackey topology is generated. Since by Mackey's Theorem 9, section 13, chapter 2, [44], $\tau(\mathcal{A}^1, \mathcal{R}^\infty)$ has the same dual as $\sigma(\mathcal{A}^1, \mathcal{R}^\infty)$, G is also bounded in τ, which implies (2.6.2)*.*
Now let $\eta = \frac{\epsilon}{L}$.

$(C3) \Rightarrow (C2)$ *Let $X_U = \Pi^{\mathrm{op}}(1_U)$ where U is a measurable set such that $\mathbb{P}(U) < \eta$. For a given $a \in \mathcal{Q}$, let $U^\pm = U \cap \{a_T^\pm - a_0^\pm > 0\}$. We have:*

$$\mathbb{E}[\pm 1_{U^\pm}(a_T^\pm - a_0^\pm)] = \langle |X_{U^\pm}|, a\rangle < \varepsilon,$$

which shows that $\mathrm{Var}(\mathcal{Q}_\pm)$ and consequently $\mathrm{Var}(\mathcal{Q})$ are uniformly integrable.

$\square$

2.6.2. Proof of Theorem 2.2.3

We postponed the proof of this theorem after the proof of Theorem 2.3.1 because we need to use its results. First of all, we show that $\overline{\rho^\#}$ is a convex and lower semi-continuous function on D_σ. Let $f, g, h \in D_\sigma$ be such that $f = \lambda g + (1-\lambda)h$ for some $\lambda \in (0,1)$. Let $b, c \in \mathcal{D}_\sigma$ be such that $\mathrm{Var}(b) = g$ and $\mathrm{Var}(c) = h$. Since $b, c \in \mathcal{D}_\sigma$ then $\mathrm{Var}(\lambda b + (1-\lambda)c) = \lambda \mathrm{Var}(b) + (1-\lambda)\mathrm{Var}(c) = f$ which gives $\lambda b + (1-\lambda)c \in \{a \in \mathcal{D}_\sigma \,|\, \mathrm{Var}(a) = f\}$. Therefore we have

$$\inf_{\{a \in \mathcal{D}_\sigma \,|\, \mathrm{Var}(a) = f\}} \rho^*(a) \leq \rho^*(\lambda b + (1-\lambda)c) \leq \lambda \rho^*(b) + (1-\lambda)\rho^*(c),$$

where in the second inequality we use the convexity of ρ^*. Taking the infimum over all b, c for which $\mathrm{Var}(b) = g$ and $\mathrm{Var}(c) = h$, we have the convexity of $\overline{\rho^\#}$.

Let $X \in \mathcal{R}^\infty$ *and define* $\phi(f) = \sup_{\{a \in \mathcal{D}_\sigma \,|\, \mathrm{Var}(a)=f\}} \langle X, a\rangle$ *for all* $f \in D_\sigma$. *We claim that* ϕ *is linear and* $\sigma(L^1, L^\infty)$*-lower semi-continuous on* D_σ. *One can consider that* $X > 1$, *since otherwise one takes* $X + \|X\|_{\mathcal{R}^\infty} + 1$ *instead.*
First, let us see that since $X = \Pi^{\mathrm{op}}(X)$ *then* $\langle X, a\rangle = \langle \Pi^{\mathrm{op}}(X), a\rangle = \langle X, \Pi^*(a)\rangle$ *for all* $a \in \hat{\mathcal{D}}_\sigma$. *This gives that*

$$\phi(X) = \sup_{\{a \in \mathcal{D}_\sigma \,|\, \mathrm{Var}(a)=f\}} \langle X, a\rangle = \sup_{\{a \in \hat{\mathcal{D}}_\sigma \,|\, \mathrm{Var}(a)=f\}} \langle X, a\rangle \ , \ \forall X \in \mathcal{R}^\infty. \tag{2.6.3}$$

Let $0 < \epsilon \leq 1$ *and define* X^ϵ *as*

$$X^\epsilon := \min\{X, X^* - \epsilon\}.$$

It is clear that X^ϵ *is a càdlàg process and then in* $\hat{\mathcal{R}}^\infty$. *Also* $(X^\epsilon)^* = X^* - \epsilon$. *Let* $\theta^\epsilon = \inf\{t \leq T \,|\, X^\epsilon = X^* - \epsilon = (X^\epsilon)^*\}$. *Thus,* θ^ϵ *is a random time and not necessarily a stopping time. It is also clear that* $X^\epsilon_{\theta^\epsilon} = (X^\epsilon)^* = X^* - \epsilon$. *Let* $f \in D_\sigma$, *and define* $\phi^\epsilon(f) = \sup_{\{a \in \hat{\mathcal{A}}^1 | \mathrm{Var}(a)=f\}} \langle X^\epsilon, a\rangle$. *Let* $a^\epsilon = (0, 1_{[\theta^\epsilon, T]} f)$. *It is clear that* $a^\epsilon \in \hat{\mathcal{D}}_\sigma$ *and* $\mathrm{Var}(a^\epsilon) = f$.
From the definition of ϕ, X^ϵ *and* θ^ϵ *we have*

$$\phi^\epsilon(f) \geq \langle X^\epsilon, a^\epsilon\rangle = \mathbb{E}[X^\epsilon_{\theta^\epsilon} f] = \mathbb{E}[(X^* - \epsilon) f].$$

On the other hand it is clear that

$$\phi^\epsilon(f) \leq \mathbb{E}[(X^\epsilon)^* f] = \mathbb{E}[(X^* - \epsilon) f].$$

This inequality along with the previous one yield $\phi^\epsilon(f) = \mathbb{E}[(X^* - \epsilon)f]$*.*
By (2.6.3) *we have*

$$\begin{aligned}
|\phi(f) - \phi^\epsilon(f)| &= \left| \sup_{\{a \in \hat{\mathcal{D}}_\sigma | \mathrm{Var}(a) = f\}} \langle X, a\rangle - \sup_{\{a \in \hat{\mathcal{D}}_\sigma | \mathrm{Var}(a) = f\}} \langle X^\epsilon, a\rangle \right| \\
&\leq \sup_{\{a \in \hat{\mathcal{D}}_\sigma | \mathrm{Var}(a) = f\}} |\langle X^\epsilon - X, a\rangle| \\
&\leq \epsilon \mathbb{E}[f] \xrightarrow{\epsilon \to 0} 0
\end{aligned}$$

This shows that $\phi(f) = \mathbb{E}[fX^*]$ *on* D_σ*. In general when* $X > 1$ *is not necessarily true, it is easy to see that* $\phi(f) = \mathbb{E}[((X + X^*)^* - X^*)f]$*, which shows that* ϕ *is linear as well as* $\sigma(L^1, L^\infty)$*-lower semi continuous. Now observe that since by Theorem 2.3.1* $\{a \in \mathcal{D}_\sigma \,|\, \mathrm{Var}(a) = f\}$ *is a convex compact set in the locally convex topological space* $(\mathcal{A}^1, \sigma(\mathcal{A}^1, \mathcal{R}^\infty))$*, a Minimax Theorem mentioned in [33] yields*

$$\overline{\rho^\#}(f) = \inf_{\mathrm{Var}(a) = f} \rho^\#(a) = \inf_{\{\mathrm{Var}(a) = f\}} \sup_{\{\rho(X) \leq 0\}} \langle -X, a\rangle = \sup_{\{\rho(X) \leq 0\}} \inf_{\{\mathrm{Var}(a) = f\}} \langle -X, a\rangle.$$

This relation along with the previous discussions yields that $\overline{\rho^\#}$ *is the supremum of a family of linear lower semi-continuous functions on* $\mathcal{D}_\sigma$*. This implies that* $\overline{\rho^\#}$ *is a lower semi-continuous function on* $\mathcal{D}_\sigma$*.*
On the other hand we have

$$\begin{aligned}
\{Y \in \mathcal{A}_{\bar{\rho}}\} &= \{\bar{\rho}(Y) \leq 0\} \\
&= \left\{ \sup_{a \in \mathcal{D}_\sigma} \{\langle -(\mathbb{E}[Y|\mathcal{F}_t])_{t \in [0,T]}, a\rangle - \rho^\#(a)\} \leq 0 \right\} \\
&= \left\{ \mathbb{E}[-Y \mathrm{Var}(a)] \leq \rho^\#(a) \,,\, \forall a \in \mathcal{D}_\sigma \right\} \\
&= \left\{ \mathbb{E}[-Yf] \leq \inf_{\{\mathrm{Var}(a) = f\}} \rho^\#(a) \,,\, \forall f \in D_\sigma \right\} \\
&= \left\{ \mathbb{E}[-fY] \leq \overline{\rho^\#}(f) \,, \forall f \in D_\sigma \right\}.
\end{aligned}$$

Now let $g \in D_\sigma$, *since* $\overline{\rho^{\#}}$ *is a convex lower semi continuous function and a supremum of linear functions on* $\mathcal{D}_\sigma$, *we have*

$$\begin{aligned}(\bar{\rho})^{\#}(g) &= \sup_{\{Y \in \mathcal{A}_{\bar{\rho}}\}} \mathbb{E}[-gY] \\ &= \sup_{\left\{\mathbb{E}[-fY] \leq \overline{\rho^{\#}}(f), \forall f \in D_\sigma\right\}} \mathbb{E}[-gY] = \overline{\rho^{\#}}(g).\end{aligned}$$

□

2.6.3. Proof of Proposition 2.3.1

We define the convex risk $\rho_1 : \hat{\mathcal{R}}^p \to \mathbb{R}$ *for* $p \neq \infty$ *by*

$$\rho_1(X) := \rho(\Pi^{\mathrm{op}}(X)).$$

It is not very difficult to see that every finite value and monotone convex function on a Banach lattice is continuous. For a proof see Proposition 3.1 [59]. Therefore, the convex risk ρ_1 *is continuous.*
On the other hand by Remark 2.2.5, $\hat{\mathcal{A}}^q$ *is the dual of* $\hat{\mathcal{R}}^p$. *Therefore by quoting the Alaoglu theorem we conclude that the set* $\{a \in \hat{\mathcal{A}}^q : \rho_1^*(a) \leq c\}$ *is* $\sigma(\hat{\mathcal{A}}^q, \hat{\mathcal{R}}^p)$-*compact for every* $c \in \mathbb{R}$. *Let us assume that* $a \in \mathcal{A}^q$. *By* (2.2.12) *we have* $\langle \Pi^{\mathrm{op}}(X), a\rangle - \rho(\Pi^{\mathrm{op}}(X)) = \langle X, a\rangle - \rho_1(X)$. *This relation implies that* $\rho_1^*(a) = \rho^*(a)$ *for* $a \in \mathcal{A}^q$, *so that* $\Pi^*\left(\{a \in \hat{\mathcal{A}}^q : \rho_1^*(a) \leq c\}\right) = \{a \in \mathcal{A}^q : \rho^*(a) \leq c\}$. *Since* $\Pi^* : \hat{\mathcal{A}}^q \to \mathcal{A}^q$ *is continuous, the set* $\{a \in \mathcal{A}^q : \rho^*(a) \leq c\}$ *is* $\sigma(\mathcal{A}^q, \mathcal{R}^p)$-*compact.*

□

2.6.4. Proof of Theorem 2.3.2

(*L*5) ⇔ (*L*6) ⇔ (*L*7) ⇔ (*L*8). *By Theorem 2.2.2 and Remark 2.2.3.*

(*L*1) ⇒ (*L*5). *Let* $(Y_n)_{n\in\mathbb{N}}$ *be a uniformly bounded sequence in* L^∞ *converging in probability to* $Y \in L^\infty$. *Since* $(Y_n)_{n\in\mathbb{N}}$ *is uniformly bounded, it is also uniformly integrable and consequently* $Y_n \to Y$ *in* L^1. *Now we have*

$$c\mathbb{P}\left(\sup_{t\leq T}|\mathbb{E}[Y_n - Y|\mathcal{F}_t]_{t\in[0,T]}| > c\right) \leq c\mathbb{P}\left(\sup_{t\leq T}\mathbb{E}[|Y_n - Y||\mathcal{F}_t] > c\right) \leq \|Y_n - Y\|_{L^1},$$

for all $c > 0$. *We used Jensen's and Doob's inequalities, in the first and second inequalities respectively. Since* $Y_n \to Y$ *in* L^1, *we have* $X_n \to X$ *in probability over* $[0,T]$, *where* $X_n = (\mathbb{E}[Y_n|\mathcal{F}_t])_{t\in[0,T]}$ *and* $X = (\mathbb{E}[Y|\mathcal{F}_t])_{t\in[0,T]}$. *From the Lebesgue property it turns out that* $\rho(X_n) \to \rho(X)$, *which by definition gives the Lebesgue property for* $\bar{\rho}$.

(*L*6) $\Rightarrow$ (*L*2). *Let* $a \in \mathcal{A}^1_+$ *be such that* $\rho^*(a) \leq c$ *for some real number* c. *By the conjugate function definition,* $\forall X \in \mathcal{R}^\infty$ *we have* $\langle X, a\rangle - \rho(X) \leq c$. *In particular, this is true for every random process like* $\Pi^{op}(Y)$ *where* $Y \in L^\infty$. *By* (2.2.13) *we conclude that* $\mathbb{E}[\mathrm{Var}(a)Y] - \bar{\rho}(Y) \leq c$ *for every* $Y \in L^\infty$. *Therefore, we have* $\mathrm{Var}(\{a \in \mathcal{A}^1_+|\rho^*(a) \leq c\}) \subseteq \{\mu \in L^1_+|\bar{\rho}^*(\mu) \leq c\}$. *By assumption* (*L*6), $\mathrm{Var}(\{a \in \mathcal{A}^1_+|\rho^*(a) \leq c\})$ *is relatively compact in* $\sigma(L^1, L^\infty)$, *hence by Theorem 2.3.1* $\{a \in \mathcal{A}^1_+|\rho^*(a) \leq c\}$ *is relatively compact in* $\sigma(\mathcal{A}^1, \mathcal{R}^\infty)$.

(*L*2) $\Rightarrow$ (*L*3) *is clear.*

(*L*3) $\Rightarrow$ (*L*1). *First we assume that* ρ *is positively homogeneous. With this assumption, for every real number* c *the set* $\{a \in \mathcal{D}_\sigma|\rho^\#(a) \leq c\}$ *is equal to* $\{a \in \mathcal{D}_\sigma|\rho^\#(a) = 0\}$. *We denote this set by* $\mathcal{Q}$.
Let X_n *be a bounded sequence in* $\mathcal{R}^\infty$ *for which for some* $X \in \mathcal{R}^\infty$, $(X_n - X)^* \xrightarrow{\mathbb{P}} 0$. *Since* ρ *is positively homogeneous (therefore sub-additive) and also decreasing we have*

$$|\rho(W) - \rho(V)| \leq \rho(-(W-V)^+) + \rho(-(V-W)^+)\ ,\ \forall W, V \in \mathcal{R}^\infty.$$

This inequality allows us to consider that $X_n \leq 0$, $X = 0$ *and* $(X_n)^* \xrightarrow{\mathbb{P}} 0$. *Using assumption* (*L*2), $\mathcal{Q}$ *is relatively compact in the topology* $\sigma(\mathcal{A}^1, \mathcal{R}^\infty)$. *Therefore, Theorem 2.3.1 gives that the closed convex set* $\mathrm{Var}(\mathcal{Q})$ *is* $\sigma(L^1, L^\infty)$-*compact and as a consequence (by Theorem 2.2.2) the convex function* $Y \mapsto \sup_{f\in\mathrm{Var}(\mathcal{Q})} \mathbb{E}[-fY]$

has the Lebesgue property. Hence by (2.2.2) *we have:*

$$0 \leq \rho(X_n) = \sup_{a \in \mathcal{Q}} \langle -X_n, a \rangle \leq \sup_{f \in \mathrm{Var}(\mathcal{Q})} \mathbb{E}[(X_n)^* f] \xrightarrow{n} 0.$$

Now consider that the convex function ρ *is not necessarily positive homogeneous. Let* X_n *and* X *be bounded in* $\mathcal{R}^\infty$ *such that* $(X_n - X)^* \xrightarrow{\mathbb{P}} 0$ *(we adopt this part of the proof from the proof of Theorem 2.4 [50]). Since* X_n *is uniformly bounded then there is a bounded sequence* $c_n \in \mathbb{R}^+$ *and a positive number* ε *such that:*

$$\rho(X_n) \leq \sup_{\rho^{\#}(a) \leq c_n} \langle -X_n, a \rangle - c_n + \varepsilon.$$

Let c *be a cluster point of* c_n *and* $I \subseteq \mathbb{N}$ *such that* $|c_n - c| < \varepsilon$ *for all* $n \in I$.
Let $\rho_1(X) := \sup_{\{\rho^{\#}(a) \leq c+\varepsilon\}} \langle -X, a \rangle$. *Since* ρ_1 *is positively homogeneous, it has the Lebesgue property. Now we have*

$$\begin{aligned}
\rho(X) &\geq \sup_{\{\rho^*(\mu) \leq c+\varepsilon\}} \langle -X, \mu \rangle - c - \varepsilon \\
&= \rho_1(X) - c - \varepsilon \\
&= \lim_{n \in I} \rho_1(X_n) - c - \varepsilon \\
&\geq \lim_{n \in I} \sup_{\{\rho^*(\mu) \leq c_n\}} \langle -X_n, \mu \rangle - c - \varepsilon \\
&\geq \lim_{n \in I} \rho(X_n) - 3\varepsilon \\
&\geq \liminf \rho(X_n) - 3\varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ *is arbitrary the proof is complete.*

(L2) $\Rightarrow$ *(L4). Let* $X \in \mathcal{R}^\infty$ *be fixed. For every* $0 < \epsilon \leq 1$, *by* (2.2.2) *there is an* $a^\epsilon \in \mathcal{D}_\sigma$ *such that* $\rho(X) \leq -\langle X, a \rangle - \rho^*(a) + \epsilon$. *Then it follows that* $\rho^*(a^\epsilon) \leq -\langle X, a^\epsilon \rangle - \rho(X) + \epsilon \leq Const(X)$, *for all* $0 < \epsilon \leq 1$, *where* $Const(X)$ *is a real number only depending on* X. *Since* ϵ *can be chosen small enough, one can see that* $\rho(X) = \sup_{\{a \in \mathcal{D}_\sigma | \rho^*(a) \leq Const(X)\}} \{-\langle X, a \rangle - \rho^*(a)\}$. *By our assumption* $\{a \in \mathcal{D}_\sigma | \rho^*(a) \leq Const(X)\}$ *is compact. Now* $\{a^{\frac{1}{n}}\}_{n \in \mathbb{N}}$ *is a sequence in* $\{a \in \mathcal{D}_\sigma | \rho^*(a) \leq Const(X)\}$, *which by compactness, has a subsequence* $\{a^{\frac{1}{n_k}}\}$ *tending to some* $a \in \{a \in \mathcal{D}_\sigma | \rho^*(a) \leq Const(X)\}$. *Taking* $\liminf$

of both sides of $\rho(X) \leq -\langle X, a^{\frac{1}{n_k}} \rangle - \rho^*(a^{\frac{1}{n}}) + \frac{1}{n}$*, by lower semi-continuity of* ρ^* *we get* $\rho(X) \leq -\langle X, a\rangle - \rho^*(a)$*. On the other hand by* (2.2.2) *we have* $\rho(X) \geq -\langle X, a\rangle - \rho^*(a)$ *which implies* $\rho(X) = -\langle X, a\rangle - \rho^*(a)$*.*

(L4) $\Rightarrow$ *(L8). Fix* $Y \in L^\infty$ *and let* $X = (\mathbb{E}[Y|\mathcal{F}_t])_{t\in[0,T]}$*. Then by assumption there is* $a^X \in \mathcal{D}_\sigma$ *in which the maximum in* (2.2.3) *is attained, i.e.* $\bar{\rho}(Y) = \rho(X) = -\langle X, a^X\rangle - \rho^\#(a^X) = \mathbb{E}[-Y\mathrm{Var}(a^X)] - \rho^\#(a^X)$*. This implies that for every* a *with* $\mathrm{Var}(a) = \mathrm{Var}(a^X)$ *we have* $\mathbb{E}[-Y\mathrm{Var}(a^X)] - \rho^\#(a^X) \geq \mathbb{E}[-Y\mathrm{Var}(a)] - \rho^\#(a)$ *and consequently* $\rho^\#(a^X) \leq \rho^*(a)\,, \forall a, \mathrm{Var}(a) = \mathrm{Var}(a^X)$*. It follows that*

$$\rho^\#(a^X) = \inf_{\{a\in\mathcal{D}_\sigma\,|\,\mathrm{Var}(a)=\mathrm{Var}(a^X)\}} \rho^\#(a),$$

which by Theorem 2.2.3 yields $\rho^\#(a^X) = \overline{\rho^\#}(\mathrm{Var}(a^X)) = (\bar{\rho})^\#(\mathrm{Var}(a^X))$*. Now it turns out that* $\bar{\rho}(Y) = \mathbb{E}[-Y\mathrm{Var}(a^X)] - \bar{\rho}^\#(\mathrm{Var}(a^X))$*, which shows* $\bar{\rho}$ *attains its maximum at* $\mathrm{Var}(a^X)$*.*

□

2.6.5. Proof of Theorem 2.3.3

($\Rightarrow$*) Is clear.*

($\Leftarrow$*) Define the convex function* ρ *by:*

$$\rho(X) := \sup_{a\in\mathcal{Q}} \langle X, a\rangle. \tag{2.6.4}$$

It is not difficult to see that $\mathrm{Var}(\mathcal{Q})$ *is convex and weakly closed. Let* $Y \in L^p$*. It is easy to see that* $\bar{\rho}(Y) = \sup_{f\in\mathrm{Var}(\mathcal{Q})} \mathbb{E}[Yf]$*. By assumption, for any* $Y \in L^p$ *there exists an* $a \in \mathcal{Q}$ *such that*

$$\rho((\mathbb{E}[Y|\mathcal{F}_t])_{0\leq t\leq T}) = \langle (\mathbb{E}[Y|\mathcal{F}_t])_{0\leq t\leq T}, a\rangle.$$

This gives $\bar{\rho}(Y) = \mathbb{E}[\mathrm{Var}(a)Y]$*. This fact with James' Theorem implies that* $\mathrm{Var}(\mathcal{Q})$ *is weakly compact. Now by Theorem 2.3.1 we deduce that* $\mathcal{Q}$ *is compact in* $\sigma(\mathcal{A}^q, \mathcal{R}^p)$*.* □

Chapter 3

RISK MEASURES ON THE SPACE OF INFINITE SEQUENCES

Résumé

Les mesures axiomatiques du risque ont été l'objet de nombreuses études et généralisations dans ces dernières années. Dans la littérature, nous trouvons principalement deux grandes écoles: les mesures cohérentes de risque [5] et les mesures de risque d'assurance [66]. Dans cet article, nous étudions une autre extension motivée par une troisième mesure axiomatique de risque qui a été introduite récemment. Dans [53], la notion de statistiques naturelles de risque, traduction libre de "natural risk statistics", est présentée comme une mesure du risque pour les bases de données, c'est-à-dire, comme une mesure du risque axiomatique définie dans l'espace $\mathbb{R}^n$*. Un inconvénient de ce type de mesures de risque est leur dépendance à l'égard de la dimension* n *de l'espace. Afin de contourner ce problème, nous proposons un moyen de définir une famille* $\{\rho_n\}_{n\geq 1}$ *de statistiques naturelles de risque dont les éléments sont définis sur* $\mathbb{R}^n$ *et liés d'une façon adéquate. Cette construction nécessite la généralisation de "natural risk statistics" dans l'espace des suites infinies* l^∞*.*

Abstract

Axiomatically based risk measures have been the object of numerous studies and generalizations in recent years. In the literature we find two main schools: coherent risk measures [5] and insurance risk measures [66]. In this note, we set to study yet another extension motivated by a third axiomatically based risk measure that has been recently introduced. In [53], the concept of natural risk statistics is discussed as a data-based risk measure, i.e. as an axiomatic risk measure defined in the space $\mathbb{R}^n$*. One drawback of these kind of risk measures is their dependence on the space dimension. In order to circumvent this problem, we propose a way to define a family* $\{\rho_n\}_{n\geq 1}$ *of natural risk statistics whose members are defined on* $\mathbb{R}^n$ *and related in an appropriate way. This construction requires the generalization of natural risk statistics to the space of infinite sequences* l^∞*.*

3.1. Introduction

Designing risk measures with the right properties is an important problem from a practical point of view and, at the same time, it leads to interesting mathematical constructions. The usual approach is to postulate some reasonable axioms and then characterize the set of risk measures that satisfy these axioms. Coherent risk measures [5] and insurance risk measures [66] are examples of such constructions.

In a recent research paper the concept of natural risk statistics has been introduced [53] in order to resolve some of the incompatibilities between these two main axiomatic risk measures. An interesting feature of this risk measure is that it is defined on $\mathbb{R}^n$*, i.e. the new risk measure assigns a value to a finite sample* $(x_1, \ldots, x_n)$*. This function measures the risk associated with a data sample from a financial (or insurance) position (no assumption on the distribution is required) instead of measuring the risk associated with the random variable itself (which requires further assumptions on the underlying distribution). One can argue that, more often than not, this is the kind of information upon which a risk manager relies to perform any risk analyzing. As a by-product, this new risk measure gives an axiomatic construction of VaR.*

This characterization is what makes natural risk statistics consistent with industrial practices. These risk measures can be found as the supremum over a set of different scenarios defined by w_i. The set containing all the scenarios $\mathcal{W}_n$ depends on n. From where we can see that a key element needed to define a risk measure in this setting is the data size (i.e. n). Different values for n lead to structurally different natural risk statistics. This inconsistency could lead two independent observers with non-disjoint collection of data of different sizes to infer substantially different risk values. This problem motivates us to define a family of natural risk statistics $\{\rho_n\}$ which are related in an appropriate way and stem from one source.

Our construction is carried out in three steps. First, we find an appropriate family of extensions $\psi_n : \mathbb{R}^n \to c_l$ or l^∞ (here l^∞ is the family of bounded sequences and c_l the set of members in l^∞ having a limit). Second, we define a suitable natural risk statistics $\rho : c_l$ or $l^\infty \to \mathbb{R}$. And finally, we combine the extension and the natural risk statistics defined on c_l or l^∞, in order to obtain a family of risk measures. The family $\{\rho_n\}_{n=1,2,\ldots}$ is defined as $\rho_n = \rho \circ \psi_n$. This procedure is illustrated in Figure 4.1.

$$
\begin{tikzcd}
c_l \textit{ or } l^\infty \arrow[r, "\rho"] & \mathbb{R} \\
\mathbb{R}^n \arrow[u, "\psi_n"] \arrow[ur, "\rho_n"'] &
\end{tikzcd}
$$

FIGURE 3.1. The commutator diagram

Through this procedure, we can construct a family of natural risk statistics that is consistently defined for data sample of all sizes. As we will see in the examples of Section 3.5, the representation of risk measures in l^∞ naturally produces families of natural risk statistics with built-in consistency.

The main goal of this note is to extend the notion of natural risk statistics to l^∞ so that we can deal with data samples of any size in a consistent way. In this setting, we suppose that we have an infinite collection of data $(x_i)_{i=1,2,...}$ that can be seen as an element in l^∞. We discuss in this work how an axiomatic risk measure ρ can be defined on l^∞.

Our motivation for studying functions on l^∞ is two-fold. First, this space allows us to study all finite collection of data without considering any bound on data size. Second extending the theory of coherent and convex risk measures to include risk measures on l^∞ is an interesting mathematical exercise on its own right.

We start with a brief discussion of the concept of natural risk statistics in Section 3.2. As we have illustrated in Figure 4.1, our construction is carried out in three steps. These different steps are the subject of subsequent sections. We discuss the problem of extending vectors from $\mathbb{R}^n$ to c_l or l^∞ in Section 3.3. The motivation behind our interest in studying functions on the subset space c_l can also be found in that section. It turns out that our interest in c_l is linked to a particular family of extensions $\psi_n : \mathbb{R}^n \to (l^\infty$ or $c_l)$, that we use in the construction described in Figure 4.1. In Section 3.4, we give the characterization of natural risk statistics in the spaces l^∞ and c_l. These results, along with the extension defined in the previous sections, will produce a family of natural risk statistics that can be used for data sets of any dimension. Finally, in Section 3.5, we illustrate this procedure with some examples and we briefly discuss some robustness features of our extension.

3.2. Natural Risk Statistics

The concept of natural risk statistics was first introduced in [53]. This notion attempts to respond to some criticized features of coherent and convex risk measures, as introduced in [5] and [41]. One criticism, recently made about coherent risk measures, is that of the absence of robustness with respect to outliers in a given data sample $(x_1, ..., x_n)$ (see for instance [29] and [53]). It turns out that there is an incompatibility between robustness and coherence for natural risk *statistics (see [29]). This fact is documented in [1] and is a consequence of the very characterization of natural risk statistics. As we will see in the last section, coherent risk measures have a representation that give more weight to larger losses and this is at the heart of this incompatibility. As it turns out, we only need to modify the subadditivity axiom in the definition of coherent risk measures (convex property for convex risk measure), in order to bring robustness features into our construction.*

In order to proceed with our discussion, we briefly present in this section some definitions and results regarding natural risk statistics. We start with the axiomatic definition of a natural risk statistics which is stated here for finite ($\mathbb{R}^n$) and infinite (l^∞ or c_l) data sets.

Definition 3.2.1. *Let $\mathbb{A}$ be either one of spaces $\mathbb{R}^n$, l^∞ or c_l. A function $\rho : \mathbb{A} \longrightarrow \mathbb{R}$ is a natural risk statistics if:*

(1) It is positive homogeneous, i.e.

$$\rho(\lambda X) = \lambda \rho(X),\ \forall\, X \in \mathbb{A}\,,$$

for any $\lambda \geq 0$.

(2) It is translation invariant, i.e.

$$\rho(X + c\mathbf{1}) = \rho(X) + c, \forall\, X \in \mathbb{A}, c \in \mathbb{R},$$

where $\mathbf{1} = (\underbrace{1, \ldots, 1}_{n-times})$ if $\mathbb{A} = \mathbb{R}^n$ and $\mathbf{1} = (1, 1, \ldots)$ if $\mathbb{A} = l^\infty$ or c_l.

(3) It is increasing, i.e.

$$\rho(X) \leq \rho(Y)\ ,$$

for all $X \leq Y$ in $\mathbb{A}$. Here, the inequality $X \leq Y$ must be understood in the component wise sense.

(4) It is comonotonic subadditive, i.e., if

$$(x_i - x_j)(y_i - y_j) \geq 0$$

for any $j \neq i$, then

$$\rho(x_1 + y_1, \ldots, x_n + y_n) \leq \rho(x_1, \ldots, x_n) + \rho(y_1, \ldots, y_n)\ ,$$

for $X, Y \in \mathbb{R}^n$, and

$$\rho(x_1 + y_1, x_2 + y_2, \dots) \leq \rho(x_1, x_2, \dots) + \rho(y_1, y_2, \dots) \; ,$$

for all $X, Y \in l^\infty$ or c_l.

(5) It is symmetric, i.e.

$$\rho(X) = \rho(X^{ij}) \; ,$$

for all $X \in \mathbb{A}$ and all $i, j > 0$. Here the sequence X^{ij} is the element in $\mathbb{A}$ which is equal component wise to X except for the i-th and j-th component which are interchanged.

Moreover, if ρ satisfies only (2), (3) and (5) we say it is a general symmetric risk measure.

We notice that if $\mathbb{A} = \mathbb{R}^n$, then Definition 3.2.1 is the one in [1] and [53]. If $\mathbb{A} = l^\infty$ or c_l, then Definition 3.2.1 is an extended definition of natural risk statistics for infinite data sets.

We also need the following definition:

Definition 3.2.2. *Let $X = (x_1, ..., x_n)$ be a vector in $\mathbb{R}^n$. We define $X^\downarrow := (x_1^\downarrow, \dots, x_n^\downarrow)$ to be the decreasing order statistics vectors of X, i.e. $x_1^\downarrow \geq \cdots \geq x_n^\downarrow$.*

We now present a representation theorem of natural risk statistics for finite data. The proof can be found in both [1] and [53]. The proof in [1] is more straightforward than the proof in [53]. Yet, the second one accepts more naturally an extension to the infinite dimension framework.

Theorem 3.2.1. *The function $\rho : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a natural risk statistics if and only if there exists a subset of weights $A \subseteq \mathbb{R}^n$ for which*

$$\rho(X) = \sup_{a \in A} \sum_{i=1}^{n} x_i^\downarrow a_i. \tag{3.2.1}$$

Furthermore, the set A in the relation (3.2.1) is convex and closed.

Remark 3.2.1. *In both [53] and [1] the authors considered the increasing order statistics $(x_{(1)}, \dots, x_{(n)})$ instead of $X^\downarrow$. This does not make any difference in the resulting theorem. We have chosen $X^\downarrow$ since this is the notation that we will use in the infinite dimensional setting.*

3.3. Extension from $\mathbb{R}^n$ to l^∞

As we have discussed, in order to proceed with our construction we need to study extensions of finite sequences into the space of infinite sequences. In this section, we give the definition of such an extension $\psi_n : \mathbb{R}^n \to l^\infty$ and some examples. These examples illustrate some features that appear when we extend the notion of natural risk statistics to l^∞. In the following, we assume that the sets c_l and l^∞ are equipped with a component wise ordering. We start with the following definition.

Definition 3.3.1. *A function $\psi_n : \mathbb{R}^n \to l^\infty$ is a natural statistics extension (or briefly extension) if*

(1) *It is component wise positive homogeneous, i.e.*

$$\psi_n(\lambda x_1, \ldots, \lambda x_n) = \lambda \psi_n(x_1, \ldots, x_n) ,$$

for any $\lambda \geq 0$.

(2) *It is component wise translation invariant, i.e.*

$$\psi_n(x_1 + c, \ldots, x_n + c) = \psi_n(x_1, \ldots, x_n) + c\,\mathbf{1} ,\ \forall c \in \mathbb{R},$$

,

(3) *It is component wise increasing, i.e. if $x_1 \geq y_1$, $x_2 \geq y_2, \ldots, x_n \geq y_n$, then*

$$\psi_n(x_1, \ldots, x_n) \geq \psi_n(y_1, \ldots, y_n) ,$$

component wise.

(4) *It is component wise comonotonic subadditive, i.e. if $(x_i - x_j)(y_i - y_j) \geq 0$ for any $j \neq i$, then the following holds component wise*

$$\psi_n(x_1 + y_1, \ldots, x_n + y_n) \leq \psi_n(x_1, \ldots, x_n) + \psi_n(y_1, \ldots, y_n) ,$$

(5) *It is symmetric, i.e.*

$$\psi_n(x_1, \ldots, x_n) = \psi_n(x_1^{ij}, \ldots, x_n^{ij}) ,$$

where the sequence $(x_1^{ij}, \ldots, x_n^{ij})$ is the element in $\mathbb{R}^n$ which is equal component wise to $(x_1, \ldots, x_n)$ except for the i-th and j-th component which are interchanged.

If we denote by $\Pi_m : l^\infty \to \mathbb{R}$ the projection on the m-th component then it is obvious that for any extension ψ_n, the family $\{\psi_n^m = \Pi_m \circ \psi_n\}$ is a family of natural risk statistics and we have the following proposition.

Proposition 3.3.1. *$\{\psi_n\}_{n\in\mathbb{N}}$ is a family of extensions if and only if, there exists a family $\{\psi_n^m\}_{n,m\in\mathbb{N}}$ of natural risk statistics for which*

$$\Pi_m \circ \psi_n = \psi_n^m \ .$$

This proposition shows that the family of extensions is as vast as the family of natural risk statistics. But we are not interested in such a large family of extensions, in this paper, we are only concerned with a somewhat smaller family. Let $\{\tilde{\rho}_n\}_{n\in\mathbb{N}}$ be a family of natural risk statistics. Then we can define the following family of extensions,

$$\psi_n(x_1,\ldots,x_n) = \left(x_1^{\downarrow},\ldots,x_n^{\downarrow},\tilde{\rho}_n(x_1,\ldots,x_n),\tilde{\rho}_n(x_1,\ldots,x_n),\ldots\right) . \tag{3.3.1}$$

As we will see in the next section, this family of extensions produces a family of natural risk statistics that only takes into account the information of data entries larger than $\tilde{\rho}_n$. This means that, regardless of the choice of $\tilde{\rho}_n$ in extension (3.3.1), the resulting risk measure $\rho_n(x_1,\ldots,x_n)$ (that commutates the diagram in Figure 4.1) is always a function of the following set,

$$\left\{x \in \{x_1,\ldots,x_n,\tilde{\rho}_n(x_1,\ldots,x_n)\} \,\middle|\, x \geq \tilde{\rho}_n(x_1,\ldots,x_n)\right\} .$$

There are a few features of this specific family of extensions that make it remarkably interesting. In particular, this extension maps any vector in $\mathbb{R}^n$ into the subspace c_l (set of members of l^∞ with a limit). This makes somewhat easier the analysis of the resulting risk measure ρ in Figure 4.1. Thus, using extension (3.3.1) in order to map things down to c_l, has at least two benefits:

(1) As we will see in Theorem 3.4.2, when working in c_l, we do not need to impose any smoothness condition on ρ,

(2) *And, as we will see in Remark 3.4.2, working in c_l, we can consider simple risk measures, like the arithmetic average. It turns out that the arithmetic average is not even well-defined in l^∞, but it is in c_l.*

The number of families of extensions that can be defined through (3.3.1) are numerous and it depends on the choice of the natural risk statistics to be used in equation (3.3.1). Examples of natural risk statistics that could be used in defining the family of extensions are $\left\{\tilde{\rho}_n = \mathrm{VaR}_\alpha\right\}$ *or* $\left\{\tilde{\rho}_n(x_1,\ldots,x_n) = \frac{x_1+\cdots+x_n}{n}\right\}$. *These choices produce risk measures that are only concerned with data entries larger than Value at Risk and the mean, respectively.*

If we want our resulting risk measure to use all information in the data set we could use $\left\{\tilde{\rho}_n(x_1,\ldots,x_n) = \min\limits_{1\le i\le n} x_i\right\}$. *In this case, all data entries larger or equal than* $\min\limits_{1\le i\le n} x_i$ *are taken into account and all data is used.*

3.4. Natural Risk Statistics on c_l and l^∞

The second ingredient in Figure 4.1 is defining a risk measure on the spaces c_l and l^∞ that can be considered as a natural extension of a natural risk statistics. In this section, we study such an extension of the concept of natural risk statistics on a larger space than the one in which it was originally defined.

Before continuing with our discussion, we would like to recall some concepts and propositions from the topology on sequences which can be found in standard texts, for instance in [2]. First c_0 is a subspace of c_l which its members have zero limit.

From Theorem 16.14 in [2], we have that the topological dual of c_l can be identified with $\mathbb{R}\oplus c_0$ under action (X,a) where this action is defined for all $a = (a_0,a_1,a_2,\ldots)\in\mathbb{R}\oplus c_0$ and $X\in c_l$ as follows

$$(X,a) = a_0x_0 + \sum_{i=1}^{\infty} a_i x_i,$$

where $x_0 = \lim\limits_i x_i$. We will also use the following simple lemma frequently in the sequel.

Lemma 3.4.1. *Let $(X_n)_{n=1,2,\ldots}$ be a bounded sequence in l^∞. Then we have:*

(1) $(X_n)_{n=1,2,...}$ *converges point wise to a member* $X \in l^\infty$ *if and only if* $weakstar - \lim X_n = X$ *in* l^∞.

(2) if also $X_n \in c_l$, $\forall n \in \mathbb{N}$ *and* $X \in c_l$, *then* X_n *converges point wise to* X *if and only if* $weak - \lim X_n = X$ *in* c_l.

PROOF. We only prove the statement for l^∞ because the same proof works in the case c_l.

First let consider that X_n be a bounded sequence, converging point wise to X. Let M be a number larger than $\|X\|_\infty$ and $\|X_n\|_\infty$, $\forall n \in \mathbb{N}$. Fix $a \in l^1$ and let $\epsilon > 0$ be an arbitrary positive number. There is $N \in \mathbb{N}$ such that $\sum_{i=N+1}^{\infty} |a_i| < \epsilon$. It is clear that

$$\left|\sum_{i=1}^{\infty} a_i(x_i^n - x_i)\right| \leq \left|\sum_{i=1}^{N} a_i(x_i^n - x_i)\right| + \epsilon M.$$

Now let $n \to \infty$. Since X_n converges point wise to X and $a \in l^1$ then

$$\limsup_n \left|\sum_{i=1}^{\infty} a_i(x_i^n - x_i)\right| \leq \epsilon M.$$

Since $\epsilon > 0$ is arbitrary then it yields that $\lim_n \left|\sum_{i=1}^{\infty} a_i(x_i^n - x_i)\right| = 0$. Since this happens for all $a \in l^1$ then we have the result.

On the other hand let X_n converges in weak star topology to X. Let $e_N \in l^1$ be a sequence which all its components are zero except the N-th component which is one. Then $|x_N - x_N^n| = \left|\sum_{i=1}^{\infty} e_i^N(x_i^n - x_i)\right| \to 0$ when $n \to \infty$. This shows that X_n converges point wise to X. □

In this paper we use the notation $\pi(X)$ *to denote* $(x_{\pi(1)}, x_{\pi(2)}...)$ *for finite or infinite vector* $X = (x_1, x_2, ...)$ *and finite permutation* $\pi \in S^n$ *for some* $n \in \mathbb{N}$, *where* S^n *denotes the set of all permutations on* $\{1, \ldots, n\}$.

In order to proceed with our construction, we need first to extend Definition 3.2.2 to the infinite sequence setting. This is, we have to define for any $X \in l^\infty$, *another element of* l^∞ *which plays the same role as* $X^\downarrow$ *in finite dimensional spaces.*

Let $X = (x_i)_{i=1,2,3,...} \in l^\infty$ *and let* s_X *be the set consisting of* $x_0 = \limsup_{i\geq 1} x_i$ *and all members of the set* $\{x_i\}_{i=1,2,3,...}$ *which are larger than* x_0. *This construction takes into account the multiplicity of entries, i.e. if we have* $N > 0$ *components* *equal to x_i (for some i) in s_X, then all N copies of x_i are in the set s_X. We now sort the elements of s_X from the largest to the smallest. We denote this sequence $X^{\downarrow}$ in order to be consistent with the notation stated in Definition 3.2.2.*

Formally, $x_1^{\downarrow} = \sup_{i\in\mathbb{N}} x_i$ the entries in $X^{\downarrow}$ are

$$x_i^{\downarrow} = \begin{cases} \text{The i-th biggest number of } s_X & \text{if } x_{i-1}^{\downarrow} > x_0 \\ x_0 & o.w. \end{cases} \tag{3.4.1}$$

We immediately notice that $\limsup_i x_i = \lim_i x_i^{\downarrow}$.
Now we have the following important lemma

Lemma 3.4.2. *For any $X \in l^{\infty}$ let π_n^X be a permutation on $\{1,...,n\}$ such that $x_{\pi_n^X(1)} \geq x_{\pi_n^X(2)} \geq \cdots \geq x_{\pi_n^X(n)}$. Then we have*

$$weak\text{-}star - \lim \pi_n^X(X) = X^{\downarrow} \ ,$$

in l^{∞} and if also $X \in c_l$ then

$$weak - \lim \pi_n^X(X) = X^{\downarrow} \ ,$$

in c_l.

Proof. Sine $\pi_n^X(X)$ and $X^{\downarrow}$ all have the same bound $\|X\|_{\infty}$, then by lemma 3.4.1 it is enough to show the point wise convergence .
It is clear that $\{x_{\pi_N^X(j)}\}_{N=1}^{\infty}$ is an increasing sequence for each fixed j. On the other hand, for any fixed $\epsilon > 0$, we know that there is infinite number of components x_i such that $x_i > x_0 - \epsilon$, where $x_0 = \limsup x_i$. Then, for every fixed j and for large enough n, we have clearly that $x_{\pi_n^X(j)} > x_0 - \epsilon$. These arguments indicate that $\lim_n x_{\pi_n^X(j)} \geq x_0$. Now we consider three cases
case one: There is a finite x_i in s_X strictly larger than x_0. Let say l is the number of these components. It is not difficult to see that there is an N such that for $n \geq N$ the set $\{x_1, \ldots, x_n\}$ contains the set $\{x_1^{\downarrow}, \ldots, x_l^{\downarrow}\}$ and hence:

$$x_{\pi_n^X(1)} = x_1^{\downarrow}, \ldots, x_{\pi_n^X(l)} = x_l^{\downarrow},$$

for $n \geq N$. This implies also $x_{\pi_n^X(k)} \leq x_0$ for $n \geq N$ and $k > l$. Now, for fixed $k > l$, by letting $n \to \infty$ one gets $x_0 \geq \lim_n x_{\pi_n^X(k)} \geq x_0$. This obviously yields

$\lim_n x_{\pi_n^X(k)} = x_0$ for every $k > l$. This completes $\lim_n x_{\pi_n^X(k)} = x_k^{\downarrow}$ for every k.

case two: There is an infinite x_i in s_X strictly larger than x_0. Let $l \in \mathbb{N}$ be fixed. Then there is a large enough N in which $\{x_1, \ldots, x_N\}$ contains $\{x_1^{\downarrow}, \ldots, x_l^{\downarrow}\}$. That obviously implies that

$$x_{\pi_n^X(l)} = x_l^{\downarrow}, \ for \ n \geq N .$$

and then $\lim_n x_{\pi_n^X(l)} = x_l^{\downarrow}$.

case three: s_X only contains x_0. This means that for all n we have $x_n \leq x_0$. Hence as we seen above since $x_0 \leq \lim_n x_{\pi_n^X}(l)$ for all l we get

$$\lim_n x_{\pi_n^X(l)} = x_0 .$$

We show then *weak-star*-limit $\pi_n^X(X) = X^{\downarrow}$ in l^∞. For c_l it is enough to observe that $(\pi_n^X(X))_0 = \limsup(\pi_n^X(X)) = x_0$, for each n.

$\square$

Functions $X \longmapsto x_i^{\downarrow}$ are examples of general symmetric risk measures. These play an important role in the characterization of a weak-star lower semi-continuous natural risk statistics.

Proposition 3.4.1. *The function $\sup^i(X) := x_i^{\downarrow}$ as a function on l^∞ is a weak-star lower semi-continuous general symmetric risk measure.*

Proof. It is clear that $X \longmapsto x_i^{\downarrow}$ satisfies conditions 2,3 and 5 of Definition 3.2.1. In order to show that $\sup^i$ is weak-star lower semi-continuous, we need to prove that the set $\{X \in l^\infty \,|\, \sup^i(X) \leq r\}$ is weak star close for each $r \in \mathbb{R}$. Since $\sup^i$ is translation invariant then it is enough to show that $F_i = \{X \in l^\infty \,|\, \sup^i(X) \leq 0\}$ is weak star close. By induction we prove that F_i's are close. For $i = 1$ it is easy since $F_1 = \{X \in l^\infty \,|\, \sup^1(X) \leq 0\} = \{X \in l^\infty \,|\, x_j \leq 0\,, \forall j\}$. So consider that $F_1, ..., F_{i-1}$ are close then we prove that F_i is close as well.

For any finite subset $C \subseteq \mathbb{N}$ let:

$$E_C := \{X \in l^\infty; x_i \geq 0, i \in C \text{ and } x_i \leq 0, i \notin C\}$$

It is easy to see that

$$F_i = F_1 \cup ... \cup F_{i-1} \cup E_i,$$

where $E_i = \cup_{|C|=i-1} E_C$.

Let X_n be a sequence in F_i and $X^n \longrightarrow X = (x_1, x_2, ...)$ in weak-star topology. If for some $1 \leq l \leq i-1$ there is a subsequence $X^{n_k} \in F_l$, then by induction hypothesis $X \in F_l \subseteq F_i$.

So unless finite members, $X^n \in E_i$. Let $C(l)$ be equal to the l-th smallest number of C i.e. $C = \{C(1), ..., C(i-1)\}$ and $C(1) < ... < C(i-1)$. We have three cases:

Case 1: There exist subsequences X^{n_k} and C^{n_k} such that $X^{n_k} \in E_{C^{n_k}}$ and $C^{n_k}(i-1)$ is bounded. Then there exist C and a sub-subsequence $X^{n_{k_m}}$ such that $X^{n_{k_m}} \in E_C$. So by closeness of E_C we get $X \in E_C \subseteq F_i$.

Case 2: There exist subsequences X^{n_k} and C^{n_k} such that $X^{n_k} \in E_{C^{n_k}}$ and $C^{n_k}(1) \longrightarrow \infty$. Then easily one can see that $\lim_k x_j^{n_k} \leq 0 \, \forall j$ and then $X \in F_1 \subseteq F_i$.

Case 3: There exist subsequences X^{n_k} and C^{n_k} such that $X^{n_k} \in E_{C^{n_k}}$ and for some $1 \leq l < i-1$, $C^{n_k}(l)$ is bounded and $C^{n_k}(l+1) \longrightarrow \infty$. Then one can find a sub-subsequence $X^{n_{k_m}}$ and a set $C' \subseteq \mathbb{N}$ such that $|C'| = l$ and $\{C^{n_{k_m}}(1), ..., C^{n_{k_m}}(l)\} = C'$. Thus $\lim_m x_j^{n_{k_m}} \leq 0$ for $j \notin C'$ and $\lim_m x_j^{n_{k_m}} \geq 0$ for $j \in C'$ which implies $X \in E_{C'} \subseteq F_l \subseteq F_i$.

$\square$

As we mentioned before, general symmetric risk measures play an important role in our discussion and we need to state one more result regarding these measures. This result is particularly interesting because it shows that general symmetric risk measures on l^∞ and c_l only take into account information from data entries larger than the $\limsup$ *of the sequence. This takes the form of the following theorem and its corollary.*

Theorem 3.4.1. *Let $\rho : l^\infty$ or $c_l \longrightarrow \mathbb{R}$ be a general symmetric risk measure. In the case l^∞ consider that ρ is lower semi-continuous with respect to weak star topology. Then $\rho(X) = \rho(X^\downarrow)$*

Proof. From Lemma 3.4.2, we know that there is a sequence of finite permutation π_n^X such that $\pi_n^X(X) \longrightarrow X^\downarrow$ in weak star topology. So by lower semi continuity of ρ we have:

$$\rho(X^{\downarrow}) \leq \liminf_{n} \rho(\pi_n^X(X)) = \rho(X). \tag{3.4.2}$$

Let $\epsilon > 0$ be an arbitrary positive number. Let N be large enough such that $x_n < x_0 + \epsilon$ for $n > N$ where $x_0 = \limsup_i x_i$. Let π_N^X be the permutation introduced earlier i.e. π_N^X is such that $x_{\pi_N^X(1)} \geq ... \geq x_{\pi_N^X(N)}$. From the definition of π_N^X, it is obvious that $x_i^{\downarrow} \geq x_{\pi_N^X(i)}$ for $i \leq N$. On the other hand $x_i^{\downarrow} + \epsilon \geq x_0 + \epsilon > x_i$ for $i > N$. So we have $\pi_N^X(X) \leq X^{\downarrow} + \epsilon\mathbf{1}$ and then by monotonicity and translation-invariance and symmetry:

$$\rho(X) = \rho(\pi_N^X(X)) \leq \rho(X^{\downarrow}) + \epsilon. \tag{3.4.3}$$

Since $\epsilon > 0$ is arbitrary from (3.4.2) and (3.4.3) we get $\rho(X) = \rho(X^{\downarrow})$. This complete the proof for l^{∞}.

Now let us consider that ρ is on c_l. Since ρ is monotone and translation invariant hence Lipschitz, it is strong continuous. It then turns out that it is strong lower semi-continuous and hence weak lower semi-continuous. Now by the same proof as above and $\lim(\pi_n^X(X)) = \lim X$ for all $n \in \mathbb{N}$, we have $\rho(X) = \rho(X^{\downarrow})$.

$\square$

Corollary 3.4.1. *There is no weak-star continuous general symmetric risk on l^{∞}.*

Proof. Without loss of generality, we can work with the normalized risk measure $\rho(X) - \rho(0)$ (one can think of this measure as the one satisfying $\rho(0) = 0$). Under this assumption along with translation invariance we have that $\rho(1, 1, \dots) = 1$.

Now, let us consider there is a general symmetric risk ρ which is weak-star continuous. Let $X = (1, 0, 1, 0, 1, 0, ...)$. Then by Theorem 3.4.1 $\rho(1, 0, 1, 0, 1, 0, ...) = \rho(1, 1, 1, ...) = 1$. On the other hand let π_n^{-X} be the permutation defined earlier for $-X$. It is obvious that $\pi_n^{-X}(X) \longrightarrow 0$ in weak star topology. Now $1 = \rho(X) = \rho(\pi_n^{-X}(X)) \longrightarrow 0$, which is a contradiction.

$\square$

Remark 3.4.1. *Contrary to what happens in $\mathbb{R}^n$, the inverse of Theorem 3.4.1 is not true anymore. For example, the function $\rho(X) = \limsup X$ is a translation invariant, symmetric and increasing function (even subadditive and positive homogeneous) but is not lower semi-continuous for weak-star topology. For example, if a sequence $X^n = (\underbrace{1, 1, 1, ..., 1}_{n-times}, 0, 0, 0, ...)$ converges to $X = (1, 1, 1, ...)$, we have $1 = \limsup_i x_i \geq \liminf_n (\limsup_i x_i^n) = 0$.*

Remark 3.4.2. *A second remark is that even the simplest example of risk measure, arithmetic average is not well defined for any member of l^∞. If we want to include measures like arithmetic average in our framework, we need to use extensions that map any vector in $\mathbb{R}^n$ into c_l, like the one defined in (3.3.1).*

We now give representation results for natural risk statistics in the the spaces c_l and l^∞. This has to be done differently for each space. We do this in two separate subsections, starting with the characterization on c_l, which poses less complications. In a second subsection we deal with the characterization on l^∞.

3.4.1. Characterization of *Natural Risk Statistics* on c_l

First we define the following subsets of c_l (or l^∞) before discussing the characterization of risk measures on c_l. Let

$$\mathcal{B} = \{X \in c_l \,|\, x_1 \geq x_2 \geq x_3 \geq ....\},$$

$$\mathcal{B}^\circ = \{X \in c_l \,|\, x_1 > x_2 > x_3 > ....\}.$$

Lemma 3.4.3. *Let ρ be a natural risk statistics on c_l. For any $Z \in \mathcal{B}^\circ$ with $\rho(Z) = 1$, there exists a vector $W = (w_0, w_1, w_2, ...)$ such that*

$$(Z, W) = 1 \tag{3.4.4}$$

$$(X, W) < 1 \;\; \forall X \in \mathcal{B} \text{ such that } \rho(X) < 1, \tag{3.4.5}$$

where $(X, W) = \sum_{i=0}^{\infty} x_i w_i$.

Proof. Let $U = \{X \in l^\infty \,|\, \rho(X) < 1\} \cap \mathcal{B}$. Since ρ is *natural risk statistics*, then U is convex and then its closure with respect to the weak topology; i.e. $\overline{U}$, is convex as well. Since ρ is translation invariant and monotone, then it is Lipschitz and then continuous in strong topology of c_l. Specially it is lower semi-continuous and then weak lower semi-continuous. This implies that $\overline{U} \subseteq \{\rho(X) \leq 1\} \cap \mathcal{B}$. On the other hand, the point Z is on the boundary of U since $\rho(Z - \epsilon 1) = 1 - \epsilon \uparrow 1$ when $\epsilon \downarrow 0$ and $\rho(Z) = 1$. So by Hahn-Banach theorem there exists a nonzero $W \in \mathbb{R} \oplus l^1$ such that,

$$(W, X) \leq (W, Z), \forall X \in \overline{U}. \tag{3.4.6}$$

□

Up to this point, we have simply followed the proof of Lemma 1 in [53]. Now, we have to adapt the proof to our setting. We can show the strict inequality happens when $X \in U$. *We can do this by contradiction. Suppose that strict inequality in (3.4.6) does not necessarily happen when* $X \in U$. *This means that there exists* $X \in U$ *such that* $(X, W) = (Z, W)$. *It is clear that*

$$(X^\alpha, W) = (Z, W), \tag{3.4.7}$$

$$\rho(X^\alpha) < 1 \ , \forall \alpha \in (0, 1), \tag{3.4.8}$$

where $X^\alpha = \alpha Z + (1 - \alpha)X$. *Since* $Z \in \mathcal{B}^\circ$ *and* $X \in \mathcal{B}$ *then* $X^\alpha \in \mathcal{B}^\circ$. *Fix some* $\alpha \in (0, 1)$ *and* $\delta > 0$. *Let* $\tilde{\epsilon}_1 = \min\{\frac{x_1^\alpha - x_2^\alpha}{3}, \delta\}$ *and* $\tilde{\epsilon}_i = \min\{\frac{x_{i-1}^\alpha - x_i^\alpha}{3}, \frac{x_i^\alpha - x_{i+1}^\alpha}{3}, \delta, \frac{1}{i}\}$ *for* $i > 1$. *Let* $\epsilon = (\epsilon_1, \epsilon_2, ...)$ *be a vector in* l^∞ *where* $\epsilon_i = sign(w_i)\tilde{\epsilon}_i$. *And finally let* $Y = X^\alpha + \epsilon$. *The vector* Y *is in* $\mathcal{B}^\circ$ *since:*

$$y_i > x_i^\alpha - \frac{x_i^\alpha - x_{i+1}^\alpha}{2} = x_{i+1}^\alpha + \frac{x_i^\alpha - x_{i+1}^\alpha}{2} > y_{i+1} \tag{3.4.9}$$

If we set δ *small enough and use relation (3.4.8) and axioms 1) through 5) in Definition 3.2.1, we get,*

$$\rho(Y) = \rho(X^\alpha + \epsilon) \leq \rho(X^\alpha + \delta 1) \leq \rho(X^\alpha) + \delta < 1. \tag{3.4.10}$$

This means that for small δ, *we have* $Y \in U$.

On the other hand, by relation (3.4.7), we have,

$$(Y,W) = (X^\alpha + \epsilon, W)$$

$$= (X^\alpha, W) + (\epsilon, W)$$

$$= (Z,W) + \sum_{i=1}^{\infty} |w_i|\tilde{\epsilon}_i > (Z,W),$$

which contradicts (3.4.6).

This finally implies that,

$$(X,W) < (Z,W), \forall X \in U. \tag{3.4.11}$$

Now, since $\rho(0) = 0$ *and then* $0 \in U$*, we have* $(Z,W) > 0$*. By rescaling* W *we get that,*

$$(Z,W) = 1 = \rho(Z).$$

This above equation along with (3.4.11) imply relation (3.4.5) and the proof is complete.

Lemma 3.4.4. *Fix* $Z \in \mathcal{B}^\circ$*. Then there is a weight* $W = (w_0, w_1, w_2, \dots)$ *such that*

$$\sum_{i=0}^{\infty} w_i = 1 \ , \tag{3.4.12}$$

$$w_i \geq 0 \ \ i = 0,1,2,3,.... \ , \tag{3.4.13}$$

$$\rho(X) \geq (X,W), \ \textit{for all } X \in \mathcal{B} \textit{ and } \rho(Z) = (Z,W). \tag{3.4.14}$$

Proof. The existence of weight W in relations (3.4.12), (3.4.14) and the fact that $w_i \geq 0$ for $i = 1,2,3,...$, follow directly from Lemma 3.4.3 and from the proof of Lemma 2 in [53]. It only remains to show that $w_0 \geq 0$.

Let $X_n = (\underbrace{1,1,1,...,1}_{n-times},0,0....)$. Then by increasing property of ρ we have,

$$1 = \rho(1) \geq \rho(X_n) \geq (X_n, W) = \sum_{i=1}^{n} w_i \ .$$

By letting $n \to \infty$, we get that $1 \geq \sum_{i=1}^{\infty} w_i$. Now, by adding w_0 to both sides in this last equation and by (3.4.12) we have,

$$1 + w_0 \geq 1,$$

which implies $w_0 \geq 0$.

$\square$

Now, we are in a position to state the characterization theorem for natural risk statistics on c_l. But before, we would like to make a few remarks regarding the proof. Our main result takes its inspiration from Theorem 1 in [53]. Our proof follows that in [53], in particular, we adapt their Lemma 1 and 2 to this new setting, which become Lemma 3.4.3 and 3.4.4, respectively. Regarding the alternative proof of Theorem 1 in [1], it cannot be used directly in our setting (see Remark 3.4.3). Thanks to Lemmas 3.4.3, 3.4.4 we can also modify those in our setting, which is given within a remark after the following theorem. Indeed, their proof strongly counts on the openness of the set $\{(x_1, ..., x_n) \in \mathbb{R}^n; x_1 > x_2 > ... > x_n\}$ while in our case the corresponding set $\mathcal{B}^\circ = \{(x_1, x_2, ...) \in c_l; x_1 > x_2 > ...\}$ is not open in strong topology. The interior of the set $\mathcal{B}$ is empty in strong and any weaker topologies. In order to see this, let $X = (x_1, x_2, \dots) \in \mathcal{B}$ and $r > 0$ be arbitrary. Denote by x_0 the $\lim x_n = \inf x_n$. There is an $N \in \mathbb{N}$ large enough such that $x_N - x_0 < \frac{r}{2}$. Let $Y = (y_1, y_2, \dots)$ be defined as $y_i = x_i$ for $i \neq N$ and $y_N = x_0 - \frac{r}{2}$. It is clear that $Y \notin \mathcal{B}$ but Y is in the open ball of radius r around X. Since X and r are arbitrary, it turns out that $\mathrm{int}(\mathcal{B})$ is empty in strong topologies on l^∞ and c_l. Obviously the interior is also empty for any weaker topology.

Theorem 3.4.2. *Let ρ be a function on c_l. The function ρ is a natural risk statistics if and only if*

$$\rho(X) = \sup_{a \in \mathcal{A}} \left\{ x_0 a_0 + \sum_{i=1}^{\infty} x_i^{\downarrow} a_i \right\}, \tag{3.4.15}$$

where $\mathcal{A}$ is a weak-star compact convex set of nonnegative sequences $a = (a_i)_{i=0,1,2,\dots}$ and $\sum_{i=0}^{\infty} a_i = 1$.

Proof. *Let us consider that ρ has a representation like (3.4.15). It is clear that ρ satisfies condition 1 through 5 in Definition 3.2.1.*

Now we prove the other direction in the implication. By Lemma 3.4.4 we know that for every vector $Z \in \mathcal{B}^{\circ}$ *there is an associated weight vector* a^Z *such that*

$$\rho(X) \geq (X, a^Z)\ ,\ \forall X \in \mathcal{B},$$
$$\rho(Z) = (Z, a^Z).$$

Now define the sets of weights A *and* $\mathcal{A}$ *as follows:*

$$A := \{a^X | \, X \in \mathcal{B}^{\circ}\},$$
$$\mathcal{A} := \overline{co}^{w^*}(A).$$

We claim that $\mathcal{A}$ *defined as above is the right choice for our statement. First of all, it is clear that for every* $a \in \mathcal{A}$ *we have* $a_i \geq 0$ *for* $i = 0, 1, 2, \ldots$ *and* $\sum_{i=0}^{\infty} a_i = 1$*. Since* $\mathcal{A}$ *is bounded, then it is obviously weak-star compact (Alaoglu's Theorem). On the other hand It is also true that* $\sup_{a \in \mathcal{A}}(X, a) = \sup_{a \in A}(X, a)\,,\, \forall X \in c_l$*. In order to see this first observe that* $\sup_{a \in \mathcal{A}}(X, a) \geq \sup_{a \in A}(X, a)\,,\, \forall X \in c_l$*. For the other inequality let* $\lambda_1, \ldots, \lambda_n$ *be nonnegative numbers suming up to one. Let* $V = \sum_{i=1}^{n} \lambda_i a^i$ *for* n *members* $a^1, \ldots, a^n$ *in* A*. Then we have* $(X, V) = \sum_{i=1}^{n} \lambda_i (X, a^i) \leq \sum_{i=1}^{n} \lambda_i \sup_{a \in A}(X, a) = \sup_{a \in A}(X, a)$*. Fix* $V \in \mathcal{A}$ *and consider the net* $V_\mu \in co(A)$ *that converges to* V *in weak star topology. We have that* $(X, V) = \lim(X, V_\mu) \leq \sup_{a \in A}(X, a)$ *and by taking supremum over* $V \in \mathcal{A}$ *we get the inequality we need. This could be done also by some polarity discussion ; see for example [32].*

Now let $X \in \mathcal{B}^{\circ}$ *be fixed. From the discussion above, we have* $\rho(X) \geq (X, a^Z)$ *for all* $Z \in \mathcal{B}^{\circ}$*, which implies that* $\rho(X) \geq \sup_{a \in A}(X, a) = \sup_{a \in \mathcal{A}}(X, a)$*. We also know that* $\rho(X) = (X, a^X)$ *and so* $\rho(X) = \sup_{a \in \mathcal{A}}(X, a)$*.*

For those $X \in \mathcal{B} \backslash \mathcal{B}^{\circ}$ *we can find a sequence* $X^n \in \mathcal{B}^{\circ}$ *such that* $\|X^n - X\|_\infty \to 0$*. Since elements of* $\mathcal{A}$ *are positive and sum up to one then, the function* $X \mapsto \sup_{a \in \mathcal{A}}(X, a)$ *is translation invariant and monotone, hence Lipschitz for* l^∞ *topology which implies*

$$\sup_{a \in \mathcal{A}}(X, a) = \lim_k \sup_{a \in \mathcal{A}}(X^k, a)\ .$$

This shows $\rho(X) = \sup_{a \in \mathcal{A}}(X, a)$ *for all* $X \in \mathcal{B}$*.*

The function ρ *is translation invariant and monotone, hence Lipschitz and consequently continuous, which imply that it is also strong and weak lower semi* *continuity. Using Theorem 3.4.1 for c_l we have that $\rho(X) = \rho(X^\downarrow)$ and since $X^\downarrow \in \mathcal{B}$ we finally obtain*

$$\rho(X) = \rho(X^\downarrow) = \sup_{a\in\mathcal{A}}(X^\downarrow, a) = \sup_{a\in\mathcal{A}}\left\{a_0x_0 + \sum_{i=1}^{\infty} a_i x_i^\downarrow\right\}.$$

Remark 3.4.3. *We would like to mention that an alternative proof following [1] could also be worked out in this setting. Here we give some hints as to how this can be achieved. Let $\hat{\rho}(X) = \rho(X) + \delta(X|\mathcal{B})$ where $\delta(.|\mathcal{B})$ is zero on $\mathcal{B}$ and $+\infty$ outside. Following the same proof in [1] we know that $\hat{\rho}$ is a convex positive homogeneous, weak lower semi-continuous function. We need to show that $\partial\hat{\rho}(X)$ is not empty for any $X \in \mathcal{B}^\circ$. Let $\rho^*(a) = \sup_{X\in c_l}\{(X, a) - \hat{\rho}(X)\}$ be the dual (or conjugate) function. Since $\hat{\rho}$ is positive homogeneous then it is easy to see that ρ^* is zero on domρ^* and $+\infty$ outside. Let a^X be the member in dual of c_l associated to X by Lemma 3.4.4. From Lemma 3.4.4, we have $\hat{\rho}(Z) \geq (Z, a^X)$ for all Z which implies that $\rho^*(a^X) \leq 0$ and hence $\rho^*(a^X) = 0$. On the other hand a^X is such that $\hat{\rho}(X) = (X, a^X)$. Now from Fenche-Moreau theorem we know that $a \in \partial\hat{\rho}(X)$ if and only if $\hat{\rho}(X) + \rho^*(a) = (X, a)$ (see [36]). It turns out that $a^X \in \partial\hat{\rho}(X)$. Following the same proof in [1] it is now clear that dom$\rho^* \subseteq \{a|\sum_{i=0}^{\infty} a_i = 1\}$. One can then show that for any member $a \in \partial\hat{\rho}(X)$, $a_i \geq 0$ for $i = 1, 2, \ldots$. Following the proof of Lemma 3.4.4 one can show that $a_0 \geq 0$. Now define $\mathcal{A} =$ dom$\rho^* \cap (\mathbb{R} \oplus l^1)^+$. The same proof as in [1] now can be carried out yielding the result.*

3.4.2. Characterization of *Natural Risk Statistics* on l^∞

As we had mentioned, extending the concept of natural risk statistics has to be done differently for each space l^∞ and c_l. In this subsection, we characterize the natural risk measures on l^∞. This representation is given in the form of the following theorem.

Theorem 3.4.3. *Let ρ be a function on l^∞. The function ρ is a weak-star lower semi-continuous natural risk statistics if and only if,*

$$\rho(X) = \sup_{a\in\mathcal{A}} \sum_{i=1}^{\infty} x_i^\downarrow a_i\ , \tag{3.4.16}$$

where $\mathcal{A}$ *is a convex set of nonnegative sequences in* l^1 *and* $\sum\limits_{i=0}^{\infty} a_i = 1$.

PROOF. If ρ has a representation like the one in (3.4.16) then, it is obvious that ρ is *natural risk statistics*. Now, let $X^n \xrightarrow{\text{weak-star}} X$, i.e. X^n converges component wise to X. So, by Proposition 3.4.1, we have $x_i^{\downarrow} \leq \liminf\limits_n x_i^{n\downarrow}$. Using Fatou lemma for a fixed $\tilde{a} \in \mathcal{A}$ we have,

$$\begin{aligned} \liminf_n \left(\sup_{a \in \mathcal{A}} \sum_{i=1}^{\infty} a_i x_i^{n\downarrow} \right) &\geq \liminf_n \sum_{i=1}^{\infty} \tilde{a}_i x_i^{n\downarrow} \\ &\geq \sum_{i=1}^{\infty} \tilde{a}_i \liminf_n x_i^{n\downarrow} \\ &\geq \sum_{i=1}^{\infty} \tilde{a}_i x_i^{\downarrow}. \end{aligned}$$

By taking supremum over $\tilde{a}$, we have finally that $\rho(X) \leq \liminf\limits_n \rho(X^n)$. This implies that ρ is lower semi-continuous, which completes the proof of the first implication.

As for the other implication, using Theorem 3.4.2 we know there exists $\tilde{\mathcal{A}}$ a weak-star compact convex set of nonnegative sequences $a = (a_i)_{i=0,1,2,\ldots}$ and $\sum\limits_{i=0}^{\infty} a_i = 1$ such that,

$$\rho|_{c_l}(X) = \sup_{a \in \tilde{\mathcal{A}}} \left\{ x_0 a_0 + \sum_{i=1}^{\infty} x_i^{\downarrow} a_i \right\} \ , \ \forall X \in c_l \ .$$

Let $\tilde{\rho}(X) = \sup\limits_{a \in \tilde{\mathcal{A}}} \sum\limits_{i=1}^{\infty} a_i x_n^{\downarrow}$ and let $X \in \mathcal{B}$ be such that $x_0 \geq 1$ and $X^n = (x_1, \ldots, x_n, 0, 0, \ldots)$. Since $x_k \geq 0 \ \forall k$, and $a_i \geq 0 \ \forall i$, then $\tilde{\rho}(X^n) \leq \tilde{\rho}(X)$. Using this along with $(X^n)_0 = 0$, $a_0 \geq 0, x_0 \geq 0$ and lower semi-continuity of ρ, we have

$$\rho(X) \leq \liminf_n \rho(X^n) = \liminf_n \tilde{\rho}(X^n) \leq \tilde{\rho}(X) \leq \rho(X) \ ,$$

which yields $\rho(X) = \tilde{\rho}(X)$. Now, let $\tilde{\mathcal{A}}_\epsilon = \{a \in \tilde{\mathcal{A}} \,|\, a_0 \leq \epsilon\}$. It is clear that $\tilde{\mathcal{A}}_\epsilon$ is increasing with respect to ϵ and also is weak star compact. So, by compactness the intersection is not empty.

Let $X \in \mathcal{B}$ be such that $x_0 \geq 1$. Since $\rho(X) = \tilde{\rho}(X)$ then, for given $\epsilon > 0$, there exists $a^\epsilon \in \tilde{\mathcal{A}}$ such that

$$\rho(X) < \sum_{i=1}^{\infty} a_i^\epsilon x_i + \epsilon \ . \tag{3.4.17}$$

On the other hand, by representation of ρ we have $\sum\limits_{i=1}^{\infty} a_i^\epsilon x_i + a_0^\epsilon x_0 \leq \rho(X)$. From these two last relations we get $a_0^\epsilon x_0 < \epsilon$ which, because of $x_0 \geq 1$, yields $a_0^\epsilon < \epsilon$. Since $\tilde{\mathcal{A}}$ is weak star compact, then there exists a net $\epsilon_k \to 0$ and $a \in \tilde{\mathcal{A}}$ such that $a^{\epsilon_k} \to a$ in weak star topology. This has two direct implications: *1)* first of all $a_0 = 0$, and *2)* since $X \in c_l$, we have $(X, a^{\epsilon_k}) \to (X, a)$.

Using (3.4.17), we have that $\rho(X) = (X, a)$ which gives

$$\rho(X) = \sup_{a \in \mathcal{A}} \sum_{i=1}^{\infty} a_i x_i \ , \tag{3.4.18}$$

where $\mathcal{A} = \cap_{\epsilon > 0} \tilde{\mathcal{A}}_\epsilon$. Notice that we can also see $\mathcal{A}$ as a subset of l^1.

Now, let $X \in l^\infty$. Since ρ is weak-star lower semi-continuous then, by Theorem 3.4.1, we have that $\rho(X) = \rho(X^\downarrow)$. Using now the fact that $\sum\limits_{i=1}^{\infty} a_i = 1$ and (3.4.18), we have

$$\begin{aligned}
\rho(X^\downarrow) - x_0 + 1 &= \rho(X^\downarrow - x_0 + 1) \\
&= \sup_{a \in \mathcal{A}} \sum_{i=1}^{\infty} a_i (x_i^\downarrow - x_0 + 1) \\
&= \sup_{a \in \mathcal{A}} \sum_{i=1}^{\infty} a_i x_i^\downarrow - x_0 + 1.
\end{aligned}$$

This completes the proof.

$\square$

This result endows us with a characterization of natural risk statistics in l^∞. This results used along with the extension defined in Section 3.3 allows us to construct a family of consistently defined natural risk statistics that can be used for data samples of any size. This will be illustrated in Section 3.5. But before doing this we would like to briefly discuss one mathematical issue regarding the characterization of risk measures in l^∞. As we have seen, it turns out that the *function limsup is important when working in the space of infinite sequences. The limsup of an infinite sequence gives the maximum trend of the infinite collection of data. Yet this simple function is not lower semi-continuous in* l^∞ *and, as such, it cannot be incorporated into our framework. This is discussed in the following subsection.*

3.4.3. A *limsup* Topology for l^∞

We start by noticing that there are simple functions that are not weak-star lower semi-continuous in l^∞*. One such function is* $\limsup$*. We have seen, that weak-star lower semi-continuous general symmetric risk measures only take into account data entries larger or equal to* $\limsup$*. But unfortunately the function* $\limsup$ *despite being convex, symmetric translation invariant and increasing (convex natural risk statistics) is not weak-star lower semi-continuous. In order to construct the smallest topology for which the function* $\limsup$ *is lower semi-continuous we should add the set* $\{X \in l^\infty | \limsup X \leq 0\}$ *and its translations to the family of the closed sets. In this subsection we carry out such a construction. We start with the following definition:*

Definition 3.4.1. *We say* X^n *converges to* X *in* $\limsup$ *convergence and write* $X^n \xrightarrow{\limsup} X$ *if* X^n *converges component wise to* X *and furthermore* $x_0 \leq \liminf\limits_n x_0^n$*.*

This convergence is clearly stronger than weak star convergence since for example $X^n = (\underbrace{1, 1, ..., 1}_{n-times}, 0, 0, ...)$ *converges in weak-star topology to* $X = (1, 1, 1, ...)$ *but, it does not converge in* $\limsup$*, i.e. we do not have* $X^n \xrightarrow{limsup} X$ *as defined in Definition 3.4.1. This convergence is also weaker than strong topology, for example* $X^n = (\underbrace{1, 1, ..., 1}_{n-times}, 1, 0, 1, 0, 1, 0, ...)$ *converges in* $\limsup$ *to* $(1, 1, 1, ...)$ *but* $\|X^n - X^{n+1}\|_{l^\infty} = 1$*. We give some remarks to relate this topology to topologies on* c_l *,* c_0*. We start by noticing that on* c_l *weak-convergence implies* $\limsup$ *convergence. More precisely, if* $(X_n)_n$ *be a sequence in* c_l *and ,*$X \in c_l$ *such that* $X^n \xrightarrow{c_l\text{-weak}} X$*, then we have* $x_k^n \to x_k$ *and* $x_0^n \to x_0$*. In turn, this implies* $X^n \xrightarrow{limsup} X$*. Clearly this implication is also true when* X_n *and* X *are in* c_0*. As* *for the reverse implication is only true in c_0 and not necessary on c_l. For instance $X^n = (\underbrace{0,0,...,0}_{n-times},1,1,...)$ tends to $(0,0,\dots)$ in* lim sup *topology but not in weak topology on c_l.*

Now we have the following theorem which gives the characterization of the natural risk statistics on l^∞ endowed with limsup topology.

Theorem 3.4.4. *The natural risk statistics $\rho : l^\infty \to \mathbb{R}$ is lower semi-continuous in limsup topology if and only if, there exists a family $\mathcal{A}$ of nonnegative sequences $\{a_i\}_{i=0,1,2,...}$ for which $\sum\limits_{i=0}^{\infty} a_i = 1$ and we have:*

$$\rho(X) = \sup_{a\in\mathcal{A}} \left\{ x_0 a_0 + \sum_{i=1}^{\infty} x_i^{\downarrow} a_i \right\}. \tag{3.4.19}$$

Proof. For the first implication, let $X^n \xrightarrow{\limsup} X$. By definition we know X^n converges component wise to X or in other words in weak-star topology. So then by Proposition 3.4.1 we have $x_i^{\downarrow} \leq \liminf\limits_n x_i^{n\downarrow}$. On the other hand, from Definition 3.4.1, we know that $x_0 \leq \liminf\limits_n x_0^n$. Now for a fixed $\tilde{a} \in \mathcal{A}$ we have,

$$\begin{aligned}
\liminf_n \left(\sup_{a\in\mathcal{A}} \sum_{i=0}^{\infty} a_i x_i^{n\downarrow} \right) &\geq \liminf_n \sum_{i=0}^{\infty} \tilde{a}_i x_i^{n\downarrow} \\
&\geq \sum_{i=0}^{\infty} \tilde{a}_i \liminf_n x_i^{n\downarrow} \\
&\geq \sum_{i=0}^{\infty} \tilde{a}_i x_i^{\downarrow} \ ,
\end{aligned}$$

where use the convention $x_0^{\downarrow} := x_0$ and $x_0^{n\downarrow} := x_0^n$. By taking supremum over $\tilde{a}$ we have $\rho(X) \leq \liminf\limits_n \rho(X^n)$.

As for the second implication, let π_n^X be the sequence of permutations defined in Section 3.4. We know that $\pi_n^X(X)$ converges component wise to $X^{\downarrow}$. On the other hand, since $\limsup X = \limsup \pi_n^X(X) = \limsup X^{\downarrow}$ we get that $\pi_n(X) \xrightarrow{\limsup} X$. Using the proof of Theorem 3.4.1, this fact yields $\rho(X) = \rho(X^{\downarrow})$.

The function $\rho|_{c_l}$ is a *natural risk statistics* on c_l, so by Theorem 3.4.2 there exists $\mathcal{A}$, a weak-star compact convex set of nonnegative sequences $a = (a_i)_{i=0,1,2,...}$ and $\sum\limits_{i=0}^{\infty} a_i = 1$ such that,

$$\rho|_{c_l}(X) = \sup_{a\in\mathcal{A}}\left\{x_0 a_0 + \sum_{i=1}^{\infty} x_i^{\downarrow} a_i\right\}. \tag{3.4.20}$$

Now, since $\rho(X) = \rho(X^{\downarrow}) = \rho|_{c_l}(X^{\downarrow})$, the proof is complete.

□

3.5. Examples of Natural Risk Statistics

In this section we give a few examples in order to illustrate how we can put together the results of the previous sections in order to construct a family of natural risk statistics through the procedure in Figure 4.1. In the following, we put together the extension defined in Section 3.3, the results in Section 3.4 and particular choices of weights in order to produce what we believe to be interesting examples. These represent only a few possible combinations of all the ingredients discussed in this paper. We would like to highlight the fact that all the examples presented here are families of natural risk statistics as originally defined in [53]. The difference here is that they have been constructed through our procedure and, as such, they are naturally derived from the representation theorems discussed in Section 3.4. Without the results developed here, these new natural risk statistics cannot be immediately identified as such. Moreover, all members of these families are consistently defined through one single set of weights that is independent of the data sample size n. This could not be achieved without a formal extension of risk measures on the spaces l^∞ and c_l.

Example 3.5.1 (Mean Exponential Risk)**.** *This example is a risk measure which combines exponential weights and an arithmetic average statistics.*

(1) Extension: We use the arithmetic average as natural risk statistics $\tilde{\rho}_n$ in extension given in (3.3.1).
Let us set $\psi_n(x_1,\dots,x_n) = \left(x_1^{\downarrow},\dots,x_n^{\downarrow}, \frac{x_1+\cdots+x_n}{n}, \frac{x_1+\cdots+x_n}{n},\dots\right)$.

(2) Weights: We use a singleton set of exponential weights in the characterization in Theorem 3.4.2, i.e. the set $\mathcal{A}$ in (3.4.15) is composed of one single infinite sequence of weights $(a_0, a_1, a_2, \dots)$ where $a_0 = 0$ and $a_i = e^{-\alpha i}(e^{\alpha}-1)$, $i = 1, 2, \dots$ for some risk parameter α.

(3) By using the characterization in Theorem 3.4.2, the resulting family of natural risk statistics is:

Let $j = j(x_1, \ldots, x_n) = \max\{i | x_i^{\downarrow} \geq \frac{x_1+\cdots+x_n}{n}\}$. *Then*

$$\rho_n(x_1, \ldots, x_n) = (e^{\alpha} - 1)^{-1} \sum_{i=1}^{j} x_i^{\downarrow} e^{-\alpha i} + \frac{e^{-\alpha j}}{e^{\alpha} - 1} \left(\frac{x_1 + \cdots + x_n}{n} \right). \tag{3.5.1}$$

Indeed, it is a straightforward exercise to verify that this risk measure has the subadditivity property (item (4) in Definition 3.2.1). Alternatively, we can notice that the sequence of exponential weights is decreasing which implies that the resulting risk measures are subadditive(see part 2 in Theorem 2 in[1]). In Figure 3.2 we illustrate the weight function of the Mean Exponential Risk.

Notice that, for every $n > 0$, *equation (3.5.1) is a natural risk statistics. This risk measure has a form that is naturally implied by the representation (3.4.15) in Theorem 3.4.2. Such a measure could not be intuitively proposed as a natural risk statistics without our construction.*

Notice that the resulting natural risk statistics in (3.5.1) is now a function of n *and of the data sample* $(x_1, x_2, \ldots, x_n)$, *hence the name statistics. It is actually a weighted sum of the sample order statistics larger or equal than the mean.*

Example 3.5.2 (Conditional Median Normal Risk)**.** *This example is a risk measure that combines weights with a normal kernel and the sample conditional median given observations larger than* VaR_{α}. *The sample conditional median is simply the sample median of those observations larger than the sample* VaR_{α}. *This statistics can itself be written as a sample VaR at a level* $\frac{1+\alpha}{2}$. *In order to see this we notice that, by the very definition of* VaR_{α}, *the proportion of observations larger than* VaR_{α} *is at most* $1 - \alpha$. *Thus, the median of these observations will be the one observation that divides this proportion in half, i.e. it will be the observation smaller or equal than the remaining* $\frac{1-\alpha}{2}$ *proportion of those observations larger than* VaR_{α}. *Clearly, this observation is itself a sample quantile (or VaR) at level* $\alpha + \frac{1-\alpha}{2} = \frac{1+\alpha}{2}$ *hence the definition* $VaR_{\frac{1+\alpha}{2}}$. *This example is a generalization of the Conditional Median Normal Risk suggested in [53].*

(1) *Extension: We use the conditional median* $\mathrm{VaR}_{\frac{1+\alpha}{2}}$, *for some level* α, *as the risk measure* $\tilde{\rho}$ *in the extension given in (3.3.1). Notice that* $\mathrm{VaR}_{\frac{1+\alpha}{2}}$ *is the conditional median of data entries larger than* VaR_α.
Let $m_\alpha = Median_\alpha(x_1, \ldots, x_n) = \mathrm{VaR}_{\frac{1+\alpha}{2}}(x_1, \ldots, x_n)$. *We set*

$$\psi_n = (x_1^{\downarrow}, \ldots, x_n^{\downarrow}, m_\alpha, m_\alpha, \ldots).$$

(2) *Weights: We use a singleton set of weights. In this example, we use a normal kernel for the components* $(a_0, a_1, a_2, \ldots)$ *of the single sequence composing the set* $\mathcal{A}$ *in the characterization in Theorem 3.4.2, i.e. let* $a_0 = 0$ *and* $a_i = \frac{1}{M} e^{\frac{|i-\mu|^2}{\sigma}}, i = 1, 2, \ldots$ *with* $M = \sum_{i=1}^{\infty} e^{\frac{|i-\mu|^2}{\sigma}}$ *for some conveniently chosen parameters* μ *and* $\sigma > 0$.

(3) *By using the characterization in Theorem 3.4.2, the resulting family of natural risk statistics is:*
Let $j = j(x_1, \ldots, x_n) = \max\left\{i | x_i^{\downarrow} \geq \mathrm{VaR}_{\frac{1+\alpha}{2}}(x_1, \ldots, x_n)\right\}$, *then*

$$\rho_n(x_1, \ldots, x_n) = \sum_{i=1}^{j} x_i^{\downarrow} \frac{1}{M} e^{\frac{|i-\mu|^2}{\sigma}} + m_\alpha \left(\frac{1}{M} \sum_{i=j+1}^{\infty} e^{\frac{|i-\mu|^2}{\sigma}} \right). \tag{3.5.2}$$

Notice that the resulting natural risk statistics in (3.5.2) is now a function of n *and of the data sample* $(x_1, x_2, \ldots, x_n)$. *It is actually a weighted sum of the sample order statistics larger or equal than the conditional median* $\mathrm{VaR}_{\frac{1+\alpha}{2}}$. *In Figure 3.3 we illustrate the weight function of the conditional median normal risk. Notice that the weight function is not decreasing and the resulting family of natural risk statistics is not coherent. Moreover, for every* $n > 0$, *equation (3.5.2) is a natural risk statistics that is naturally implied by the representation (3.4.15) in Theorem 3.4.2. Such a measure could not be intuitively proposed as a natural risk statistics without our construction.*

Example 3.5.3 (Multi-Conditional Median Normal Risk)**.** *This example extend the idea of the previous example by considering a larger set of weight sequences* $\mathcal{A}$ *in the representation (3.4.15) of Theorem 3.4.2. We do this, by considering all possible means* $N \in \mathbb{N}$ *for the parameter* μ *in our normal kernel. The result is a generalization of the previously defined Conditional Median Normal Risk.*

(1) *Extension: Like in the previous example, we use the conditional median* $\mathrm{VaR}_{\frac{1+\alpha}{2}}$, *for some level* α, *as the risk measure* $\tilde{\rho}$ *in the extension given in (3.3.1).*

Let $m_\alpha = Median_\alpha(x_1, \dots, x_n) = \mathrm{VaR}_{\frac{1+\alpha}{2}}(x_1, \dots, x_n)$.*Define*

$$\psi_n = (x_1^\downarrow, \dots, x_n^\downarrow, m_\alpha, m_\alpha, \dots).$$

(2) *Weights: As a set* $\mathcal{A}$, *in the characterization in Theorem 3.4.2, we use weight sequences with normal-based entries. In other words, we consider all possible normal kernels for the components* $\{(a_{0,N}, a_{1,N}, a_{2,N}, \dots)\}_{N\in\mathbb{N}}$ *of sequences in* $\mathcal{A}$, *i.e. for* $N \in \mathbb{N}$, *let* $a_0 = 0$ *and*$a_{i,N} = \frac{1}{M_{N,\sigma}} e^{\frac{|i-N|^2}{\sigma}}$ *with* $M_{N,\sigma} = \sum_{i=1}^{\infty} e^{\frac{|i-N|^2}{\sigma}}$ *for some conveniently chosen parameter* $\sigma > 0$.

(3) *By using the characterization in Theorem 3.4.2, the resulting family of natural risk statistics is:*

Let $j = j(x_1, \dots, x_n) = \max\left\{i | x_i^\downarrow \geq \mathrm{VaR}_{\frac{1+\alpha}{2}}(x_1, \dots, x_n)\right\}$. *Then*

$$\rho_n(x_1, \dots, x_n) = \sup_{N\in\mathbb{N}} \left(\sum_{i=1}^{j} x_i^\downarrow \frac{1}{M_{\sigma,N}} e^{\frac{|i-N|^2}{\sigma}} + m_\alpha \left(\frac{1}{M_{\sigma,N}} \sum_{i=j+1}^{\infty} e^{\frac{|i-N|^2}{\sigma}} \right) \right). \quad (3.5.3)$$

Notice that the resulting natural risk statistics in (3.5.3) is now a function of n *and of the data sample* $(x_1, x_2, \dots, x_n)$. *It is actually the supremum, over all possible values of the parameter* μ, *of weighted sums of the sample order statistics larger or equal than the conditional median* $\mathrm{VaR}_{\frac{1+\alpha}{2}}$. *In Figure 3.4 we illustrate the weight function of the multiple conditional median normal risk. Moreover, for every* $n > 0$, *equation (3.5.2) is a natural risk statistics that is naturally implied by the representation (3.4.15) in Theorem 3.4.2. Such a measure could not be intuitively proposed as a natural risk statistics without our construction.*

3.5.1. Robustness Properties

In this section, we denote by $\mathcal{D}_p$ *the space of distributions with finite p-th moment. Now let* $\rho : \mathcal{D}_p \to \mathbb{R}$ *be a distribution-based risk measure and let* $\widehat{\rho_n} : \mathbb{R}^n \to \mathbb{R}$ *be a historical risk estimator.*

The risk measures discussed in this paper are functions of data samples (hence the name statistics). In the last two decades there have been numerous studies *on robustness properties of data statistics (see [47] and references therein). In particular, we can study the robustness of our examples within the framework laid out in [29]. This is, if $\rho(F_X)$ is a distribution-based risk measure, we say that a historical estimator of this measure $\widehat{\rho}_n(x_1, \ldots, x_n)$ is robust if a small variation from the distribution F_X results in a small change in the distribution of the estimator. In order to formally state this we need the following notation. We denote by $\mathcal{L}_n(\widehat{\rho}_n, F)$ the law of $\widehat{\rho}_n(x_1, \ldots, x_n)$ where $x_1, \ldots, x_n$ is a random sample of size $n \geq 1$ from F. Moreover, let d_P denote the Prohorov metric for probability measures. A formal definition of robustness can now be given.*

Definition 3.5.1. *We say that the historical estimator $\widehat{\rho}_n$ is $\mathcal{D}_p$-robust at F if, for any $\epsilon > 0$, there exist $\delta > 0$ and $n_0 \geq 1$ such that if $G \in \mathcal{D}_p$ and $d_P(F, G) < \delta$ then $d_P(\mathcal{L}_n(\widehat{\rho}_n, F), \mathcal{L}_n(\widehat{\rho}_n, G)) < \epsilon$ for all $n > n_0$.*

Now, let us consider distribution-based risk measures $\rho_\phi : \mathcal{D}_p \to \mathbb{R}$ of the form

$$\rho_\phi(F) = \int_0^1 VaR_u(F)\, \phi(u)\, du\ , \quad F \in \mathcal{D}_p\ , \tag{3.5.4}$$

where ϕ is a density function in $L^q(0,1)$ such that $\frac{1}{p} + \frac{1}{q} = 1$. It turns out that for this type of distribution-based risk measures of the form (3.5.4), we have the following very interesting result given in [29],

Proposition 3.5.1 (Corollary 2 in [29])**.**

(1) Historical estimators of distribution-based risk measures of the form (3.5.4) which are coherent, i.e. with decreasing weighting function ϕ, are not $\mathcal{D}_p$-robust at any F for $\frac{1}{p} + \frac{1}{q} = 1$.

(2) For a $F \in \mathcal{D}_p$ such that no discontinuity of ϕ coincides with a discontinuity of the quantile function of F, the historical estimator of a distribution-based risk measure of the form (3.5.4) is $\mathcal{D}_p$-robust at F if and only if supp(ϕ) $\in [\bar{\beta}, 1 - \bar{\beta}]$, for some $\bar{\beta} > 0$.

In other words, an estimator of a distribution-based risk measure of the form (3.5.4) is not robust if the weighting function is decreasing. Moreover, the robustness of the estimator of such a distribution-based risk measure depends on the support of the weighting density ϕ in the representation (3.5.4). If this support *takes the form of a closed interval which is strictly contained within* $[0,1]$*, then the corresponding risk measure estimator is robust.*

Proposition 3.5.1 has interesting implications for some of our examples. In order to see how the natural risk statistics in our first two examples are historical estimator of distribution-based risk measure of the form in (3.5.4), let us consider the following distribution-based risk measure,

$$\begin{aligned} \rho(F) &= \int_0^1 (VaR_u(F) \vee VaR_\alpha(F))\ \phi(u)\,du \\ &= VaR_\alpha(F) \int_0^\alpha \phi(u)\,du + \int_\alpha^1 VaR_u(F)\,\phi(u)\,du\ , \end{aligned} \tag{3.5.5}$$

where $\phi : [0,1] \to \mathbb{R}^+ \cup \{\infty\}$ *is a weight function, i.e.* $\int_0^1 \phi(u)\,du = 1$*.*

We notice that the risk measure in (3.5.5) is of the form (3.5.4) and, as such, Proposition 3.5.1 would apply for their estimators. In order to write (3.5.5) in the form (3.5.4), let us define the following weight function

$$\tilde{\phi}(x) = \begin{cases} \phi(s)\ , & \alpha < s \leq 1\ , \\ \\ (\int_0^\alpha \phi(t)\,dt)\delta_\alpha\ , & s = \alpha\ , \\ \\ 0 & 0 \leq s < \alpha\ . \end{cases} \tag{3.5.6}$$

We can now write the risk measure in equation (3.5.5) in the form (3.5.4) as follows,

$$\rho(F) = \int_0^1 VaR_u\,\tilde{\phi}(u)\,du\ , \tag{3.5.7}$$

where $\tilde{\phi}$ *is the well-defined weight function given in (3.5.6).*

A first remark regarding measures of the form in (3.5.5) is that they have an alternative form for special cases, in terms of a random variable with distribution F_X*. Let us consider for a moment that* F_X *has an inverse. Then* $\mathrm{VaR}_\alpha(X) = F_X^{-1}(\alpha)$*, and by a simple change of variable* $u = F_X(y)$ *in (3.5.5), we get easily the following equivalent form:*

$$\rho(X) = \mathbb{E}\left[(X \vee VaR_\alpha(F_X))\,\phi\left(F_X(X)\right)\right]\ . \tag{3.5.8}$$

We find this form particularly informative in terms of the interpretation for such risk measure. It is the expectation of the weighted values larger than VaR_α, where the weights are given as a function of the probability of observing such large values.

Now, let $(x_1, \dots, x_n)$ be a random sample of a continuous distribution function F. We can now construct the following empirical distribution,

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{\{x_i \le x\}} , \qquad x \in \mathbb{R} . \tag{3.5.9}$$

It is well-known that this empirical distribution is a sample-based functional estimator of F. In order to see how our first two examples of natural risk statistics are historical estimators of some distribution-based risk measure like that in (3.5.5), let us define the following alternative measure defined through a distortion of the underlying distribution,

$$\tilde{\rho}(F) := \rho(k_n \circ F) , \tag{3.5.10}$$

where $k_n : [0,1] \to [0,1]$ is continuous increasing piece-wise linear function connecting $(\frac{i}{n}, \frac{2^i-1}{2^i})$ to $(\frac{i+1}{n}, \frac{2^{i+1}-1}{2^{i+1}})$, for $0 \le i \le n-2$, and connecting $(\frac{n-1}{n}, \frac{2^{n-1}-1}{2^{n-1}})$ to the point $(1,1)$. We denote each interval over which k_n is linear, with $I_1, \dots I_n$, and the restriction $k_n|_{I_i}$ with l_i. It is clear that $l_i(x) = c_i(x - x_i) + b_i$ for some $c_i > 0$, $b_i, x_i \ge 0$ where $i = 1, \dots, n$.

Here, the function k_n plays the role of an auxiliary transformation function that serves as a distortion. In fact, it is clear that the function $k_n \circ F_n$ is the probability distribution of the following random variable,

$$\widetilde{X}(\omega) = \sum_{i=1}^{n-1} x_i^{\downarrow} \mathbb{I}_{(\frac{1}{2^i}, \frac{1}{2^{i-1}}]}(\omega) + x_n^{\downarrow} \mathbb{I}_{[0, \frac{1}{2^{n-1}}]}(\omega) , \quad \omega \in \Omega = [0,1] , \tag{3.5.11}$$

where $(x_1^{\downarrow}, \dots, x_n^{\downarrow})$ is the vector of decreasing order statistics of the random sample.

We can construct estimators of the risk measure in (3.5.5). In order to construct an estimator $\tilde{\rho}$ through (3.5.10), we only need a particular choice for the weight function ϕ. For instance, if we use the following weight function

$$\phi(\omega) = \sum_{i=1}^{\infty} 2^i a_i \mathbb{I}_{(\frac{1}{2^i}, \frac{1}{2^{i-1}}]}(\omega) , \quad \omega \in \Omega = [0,1] , \tag{3.5.12}$$

in the expression for the risk measure (3.5.5), we have

$$\tilde{\rho}(F_n) = \sum_{i=1}^{j} a_i x_i^{\downarrow} + \left(\sum_{i=j+1}^{\infty} a_i \right) x_j^{\downarrow} \,, \tag{3.5.13}$$

where $j = \max\{i \mid x_i^{\downarrow} \geq VaR_\alpha(x_1, \ldots, x_n)\}$*. If we compare equation (3.5.13) with (3.5.1) and (3.5.2) we can see that they have the same form. In fact, by an appropriate choice of weights* $(a_1, a_2, \ldots)$*, we can recuperate our first two examples. In light of this, we can study the robustness properties of (3.5.13) through the distorted measure (3.5.10) using Proposition 3.5.1. All we need to show is that* (3.5.10) *has the same form as* (3.5.4)*.*

For any continuous cumulative distribution function F*, we have,*

$$\begin{aligned}
\tilde{\rho}(F) &= \rho(k_n \circ F) \\
&= \int_0^1 \mathrm{VaR}_u(k_n \circ F)\tilde{\phi}(u)du \\
&= \int_0^1 (k_n \circ F)^{-1}(u)\tilde{\phi}(u)du \\
&= \int_0^1 F^{-1}(k_n^{-1})(u)\tilde{\phi}(u)du \\
&= \sum_{i=1}^{n} \int_{I_i} F^{-1}(l_i^{-1}(u))\tilde{\phi}(u)du \\
&= \sum_{i=1}^{n} \int_{l_i^{-1}(I_i)} F^{-1}(y)\tilde{\phi}(l_i(y))c_i dy \\
&= \int_0^1 F^{-1}(y)\tilde{\tilde{\phi}}(y)dy \\
&= \int_0^1 \mathrm{VaR}_u\tilde{\tilde{\phi}}(u)du,
\end{aligned}$$

where $\tilde{\tilde{\phi}}(y) = \sum_{i=1}^{n} c_i 1_{l^{-1}(I_i)}(y)\tilde{\phi}(l_i(y))$*. Therefore* $\tilde{\rho}(F)$ *has the same form as* (3.5.4) *and we can use Proposition 3.5.1 to study the robustness of our natural risk statistics in (3.5.1) and (3.5.2).*

Now we can see that our first family of natural risk statistics, the so called mean exponential risk, can be seen as a historical estimator of a distribution-based risk measure when we choose exponential weights $a_i = e^{-\alpha i}(e^\alpha - 1)$ *in (3.5.12). Since these weights are decreasing, we have by Proposition 3.5.1 that the mean* *exponential risk statistics is not a robust. This is particularly interesting because illustrates the incompatibility of coherence and robustness in a single risk measure. And we can intuitively understand the mechanics behind this fact. If we want a risk measure to be coherent then the weights in (3.5.12) have to be decreasing, but if the weights are decreasing this means that we are giving more weight to the largest observations, hence yielding estimators that are more sensitive to sample outliers.*

As for the second example of natural risk statistics, the so-called conditional median normal risk, we can see that equation (3.5.13) is the conditional median normal risk statistics, if we use a normal kernel for the weights $(a_1, a_2, \dots)$. *This is, if we use the weights* $a_i = \frac{1}{M} e^{\frac{|i-\mu|^2}{\sigma}}$ *with* $M = \sum_{i=1}^{\infty} e^{\frac{|i-\mu|^2}{\sigma}}$ *and for some conveniently chosen parameters* μ *and* $\sigma > 0$, *then the conditional median normal risk can be seen as a historical estimator of (3.5.5). In view of Proposition 3.5.1, we can see that the robustness of the conditional median normal risk statistics depends on the support of the weight function* $\tilde{\phi}$. *In particular, it is clear that supp*$\tilde{\phi} = [\alpha, 1]$ *and so the robustness of the estimator only depends on the right end of the support of* ϕ. *There are many ways of defining a weight function* ϕ, *for instance, one could envision a definition that would have a right end of its support away from one, guaranteeing the robustness of the estimator (3.5.13).*

One final remark regarding the risk measures of the form (3.5.5). This particular form is suggested by the structure of natural risk statistics as produced by our construction. We notice that, for risk measures in c_l *and* l^{∞}, *the entries smaller than the limsup of the sequence are not taken into account (see the proof of Theorem 3.4.1). This fact brings about the idea of considering risk measures, like (3.5.5) in the first place. These measures only take into account data entries larger than a conveniently chosen quantity, like* VaR_{α} *for example. We believe that these risk measures deserve further analysis.*

Acknowledgments

We would like to thank an anonymous referee for a thorough reading of our manuscript and very useful suggestions and comments that have greatly improved the presentation of this contribution.

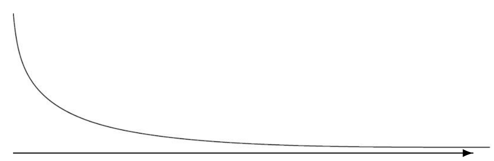

Figure 3.2. Weights for Mean Exponential Risk

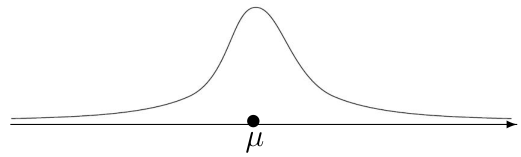

FIGURE 3.3. Weights for the Conditional Median Normal Risk

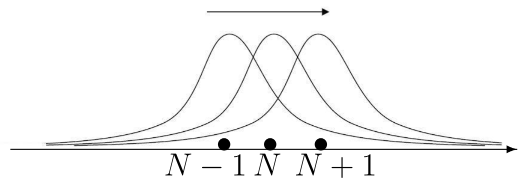

FIGURE 3.4. Weights for the Multi-Conditional-Median Normal

Chapter 4

GOOD DEALS AND THE COMPATIBLE MODIFICATION OF RISK AND PRICING RULES: A REGULATORY TREATMENT

Résumé

Dans cet article nous étudions la situation dans laquelle un marché peut être déstabilisé en présence de "bonnes affaires" (Good Deals). Une bonne affaire est une situation financière à coût zéro qui n'entraîne aucun risque. Nous étudions les bonnes affaires dans un scénario où les entreprises utilisent des mesures cohérentes pour évaluer leurs risques et où les prix du marché sont déterminés par une règle de tarification sous-linéaire. Le résultat le plus important dans ce travail est l'observation que l'existence d'une bonne affaire est équivalente à l'incompatibilité entre la règle de tarification et la mesure de risque. L'incompatibilité a été introduite et étudiée dans [13]. Nous nous penchons sur cette situation du point de vue réglementaire afin d'exclure de bonnes affaires avec l'intention de stabiliser les marchés financiers. Nous proposons quelques façons pratiques de modifier une mesure du risque de telle sorte qu'un organisme régulateur puisse établir des niveaux appropriés de capital pour les institutions financières.

Abstract

In this paper we study a situation in which a market might be destabilized in the presence of Good Deals. A Good Deal is a zero-cost financial position that

does not produce any risk. We study Good Deals in a scenario whereby a firm uses decision-making tools based on a coherent risk measure, and where the market price is determined with a sub-linear pricing rule. The most important observation of this work is that the existence of a Good Deal is equivalent to the incompatibility between the pricing rule and the risk measure. Incompatibility has been introduced and studied in [13]. In this paper, we look at this situation from a regulatory point of view in order to rule out Good Deals, with the purpose of stabilizing financial markets. We propose some practical ways of modifying a risk measure such that a regulator can set appropriate levels of capital requirements for financial institutions, in order to be considered in a safe position.

4.1. Introduction

Stability of financial markets is one of the biggest concerns of regulators, in particular central banks. In the last century the world has witnessed many financial crises that have provoked regulators to establish some rules in order to make markets safer and more stable. For example, in the European Union, Basle II (finance) and Solvency II (insurance) contain sets of rules which the industry section should respect in order to place corporations in a safer position. Following these rules, any corporation computes its "capital reserves", i.e. additional capital devoted mainly to overcome periods of loss in their economic activities. The appropriate size of reserve could be considered as the risk level associated with the firm's activities. The importance of these rules, and accordingly the "capital reserves" is to keep the markets in a safer and more stable state. It is generally accepted that stability of a market is mainly reached while the market is in equilibrium. The general theory of market equilibrium has been developed during the last century (see [30]). It is also known that equilibrium balances the market participant's needs and their preferences. In general, the state of stability is an outcome of a fair allocation of available resources among market participants. However, one cannot always rely on the existence of an equilibrium while there are financial opportunities which destabilize a market. Most of the time, market destabilizers are financial positions deemed to be simultaneously safe and profitable. The best *known example of such positions are arbitrages. An arbitrage is easily detectable and cannot survive for a long time in a market. But Arbitrages are not the only positions which destabilize a market.*

*In recent years, different risk measures have been used in financial institutions and regulatory sectors in order to assess the risk of financial positions, and in order to calculate the capital requirement. Sometimes, these risk measures provoke a new generation of market destabilizers. These financial positions are the major objectives we will study in this paper. We study a kind of pathological financial position called a Good Deal. These kinds of positions are introduced and studied in [28] and [19]. Cochrane and Saà-Requejo [28] first introduce the notion of a Good Deal as a financial position with particularly high Sharpe ratio. In that work, the authors assume that Good Deals do not exist in market equilibrium, and they show that this assumption holds if and only if there is a bound on the variance of the members of the Stochastic Discount Factor set (SDF). In Cochrane and Saà-Requejo [28], this problem is analyzed for the one-period, multi-period, and the continuous time settings. The definition of a Good Deal has been extended in Černý and Hodges [19], where the authors define a set of "desirable" positions. They define a Good Deal as a desirable position with the non-negative price and use the No Good Deal assumption to price the claims in an incomplete market. In another work, Černý [**?**] defines a Good Deal by mean of a generalized Sharpe ratio, developing the ideas in Cochrane and Saà-Requejo [28] and Černý and Hodges [19].*

*Björk and Slinkor [**?**] extend the results of Cochrane and Saà-Requejo [28] to a dynamic setting with a general Markov process, allowing a study of the Good Deal bounds for processes with jumps. Cherny [25] extends the definition of Good Deals to positions with a high performance ratio (a generalization of Sharpe ratio). All these works aim to price a financial position in an incomplete market when equilibrium is reached.*

Our work differs from the existing literature in two ways. First, we use our results for regulatory purposes (assessing the capital requirement), not for pricing. Second, we are interested in investigating a situation in which a Good Deal *exists. We found that underestimating the risk of a financial position (or under-capitalization) produces a Good Deal. To avoid under-capitalization, the financial institutions have to modify their risk measures to ones which always dominate the primary risk measures in use. In fact, the modified risk measure must dominate the primary risk measure in addition to dominating the short selling price. In this paper, we propose two ways of modifying a risk measure. The first regards the fundamentals of the risk user, and the second regards the fundamentals of the market. We also focus on concrete risk measures. Special attention is devoted to CVaR because this coherent risk measure has become very popular among researchers and practitioners. We apply our findings to CVaR so as to build the Compatible Conditional Value at Risk (CCVaR) in a general incomplete market. In an incomplete market, Compatible CVaR can be found by seeking a stochastic discount factor with the smallest European call option price. This modifies the discussion in Balbás and Balbás [13], in which Compatible CVaR is introduced in a complete perfect market (i.e. SDF is a singleton).*

This paper is organized as follows. In Section 2, we will present the notations and the general framework with which we will work. We will consider an Arbitrage-free market (in general incomplete and/or imperfect) with a sub-linear pricing rule π *and a coherent risk measure* ρ*. In Section 3, we define the concept of a Good Deal, inspired by definitions in Černý and Hodges [19] and Cherny [25]. We will show that incompatibility is equivalent to the existence of Good Deals. In Section 4, we will show the existence of a minimal compatible modification of a coherent risk measure. We will see that the existence of a minimal compatible modification is tied to the existence of a minimal point of a partial order on SDF. In Section 5, which constitutes the second part of the paper, we will propose two ways of modifying a risk measure.*

4.2. Preliminaries and Notation

Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ *composed of the set* Ω *representing the "states of the world", a* σ*-field* $\mathcal{F}$ *and a probability measure* $\mathbb{P}$*. Let* $p, q \in [1, \infty]$

be two numbers such that $1/p+1/q=1$. *For* $p \neq \infty$, L^p *denotes the space of real-valued random variables* X *on* Ω *such that* $\mathbb{E}\left(|X|^p\right) < \infty$ *where* $\mathbb{E}$ *represents the mathematical expectation. The space* L^∞ *consists of all bounded random variables. Recall that according to the Riesz Representation Theorem,* L^q *is the dual space of* L^p *when* $p \neq 1, \infty$. *We mainly endow the space* L^p *and* L^q *by two topologies, first the norm topology and second the topology induced by* L^q *i.e. the coarsest topology in which all members of* L^q *are continuous. As usual the latter topology is called by weak topology and is denoted by* $\sigma(L^p, L^q)$ *(there is one exception for* $p=\infty$ *when* $\sigma(L^\infty, L^1)$ *is called weak star topology).*

In this paper we consider only two periods of time, today and tomorrow, represented by 0 *and* T *respectively. Every random variable represents the pay-off of a financial position at time* T. *Whenever we talk about risk or price of a financial position we mean the present value of the price and the present risk associated to the financial position. In addition, to simplify the discussions we consider that the interest rate is zero.*

Let us assume that $\mathcal{X} \subset L^p$ *is a closed convex cone containing* $\mathbb{R}$ *(the set of real numbers), representing all viable pay-offs, i.e. for every* $X \in \mathcal{X}$ *there is a price associated with* X.

Definition 4.2.1. *A* L^p*-continuous mapping* $\pi : \mathcal{X} \to \mathbb{R}$ *is a sub-linear pricing rule if*

i) $\pi(X+k) = \pi(X) + k, \forall X \in \mathcal{X}, \forall k \in \mathbb{R}$;

ii) $\pi(\lambda X) = \lambda \pi(X), \forall X \in \mathcal{X}, \forall \lambda > 0$;

iii) $\pi(X+Y) \leq \pi(X) + \pi(Y), \forall X, Y \in \mathcal{X}$;

iv) $\pi(X) \leq \pi(Y), \forall X, Y \in L^p$ *and* $X \leq Y$.

Remark 4.2.1. *The pricing rule* π *can be for example considered the super-replication price, when* $\mathcal{X}$ *consists of all random variables like* X *such that there exists a viable self-financing process which can super-hedge* X.

Definition 4.2.2. *A continuous mapping* $\rho : L^p \to \mathbb{R}$ *is a coherent risk measure if*

1) $\rho\left(X+k\right) = \rho\left(X\right) - k$ *for every* $X \in L^p$ *and* $k \in \mathbb{R}$;

2) $\rho\left(\lambda X\right) = \lambda \rho\left(X\right)$ *for every* $X \in L^p$ *and* $\lambda > 0$;

3) $\rho(X+Y) \leq \rho(X) + \rho(Y)$ for every $X, Y \in L^p$;

4) $\rho(X) \leq \rho(Y)$ for every $X, Y \in L^p$ and $X \geq Y$.

A particularly interesting example is the Conditional Value at Risk (CVaR) of Rockafellar et al. [58] defiend as

$$\mathrm{CVaR}_\alpha(X) = \frac{1}{\alpha}\int_0^\alpha \mathrm{VaR}_s(X)ds.$$

Let

$$\Delta_\rho := \{Z \in L^q | -\mathbb{E}(XZ) \leq \rho(X), \forall X \in L^p\}. \tag{4.2.1}$$

The set Δ_ρ is obviously convex. Bearing in mind the Representation Theorem 2.4.9 *in Zalinescu [67] for $p \neq \infty$, and using a proof similar to that of the Representation Theorem of a risk measure, from what is stated in Rockafellar et al. [58], it can be seen that Δ_ρ is $\sigma(L^q, L^p)-$compact, and*

$$\rho(X) = \max_{Z \in \Delta_\rho} \mathbb{E}[-ZX] \ , \ \forall X \in L^p. \tag{4.2.2}$$

Furthermore, by 1) *and* 4) *of Definition 4.2.2 one can see that*

$$\Delta_\rho \subset \{Z \in L^q_+ | \mathbb{E}(Z) = 1\}. \tag{4.2.3}$$

By means of the Hahn-Banach Separation Theorem, one can easily prove that if $\Delta_\rho \subset L^q$ is convex and $\sigma(L^q, L^p)-$compact, and Δ_ρ satisfies (4.2.3)*, then there exists a unique continuous mapping $\rho : L^p \to \mathbb{R}$ satisfying* 1), 2), 3) *and* 4) *such that* (4.2.2) *holds.*

For $p = \infty$, in order to have the same representation, ρ needs to have the Fatou property introduced by Delbaen [33]. We say that ρ has the Fatou property if for any bounded sequence $\{X_n\}_n \subseteq L^\infty$, converging in probability to X we have that $\rho(X) \leq \liminf_n \rho(X_n)$. For coherent risk measures this is equivalent to the continuity from above i.e., for every sequence $\{X_n\}_n$ in L^∞ such that $X_n \downarrow X$ we have that $\rho(X_n) \to \rho(X)$ (see Delbaen [33]). With this assumption Δ_ρ is a subset of L^1, but not in general $\sigma(L^1, L^\infty)$-compact. In the sequel for $p = \infty$ we also add the assumption that Δ_ρ is $\sigma(L^1, L^\infty)$-compact, which with the aid of the Dunford-Pettis Theorem means that Δ_ρ is uniformly integrable. It is worth mentioning that the $\sigma(L^1, L^\infty)$-compactness is equivalent to the so-called Lebesgue *property of ρ defined in Jouini et al. [50]. A coherent risk measure ρ satisfies the Lebesgue property if for any bounded sequence $\{X_n\}_n \subseteq L^\infty$ converging in probability to X we have that $\rho(X_n) \to \rho(X)$. For coherent risk measures this property is also equivalent to the continuity from below i.e. if $X_n \uparrow X$ then $\rho(X_n) \to \rho(X)$. For further discussions see for example Föllmer and Schied [43] Proposition 4.21. It is also important to know that most common law invariant coherent (convex in general) risk measures display this property. For instance, for the coherent risk measure CVaR_α (where $\alpha \in (0,1)$ is a confidence level) we know that $\Delta_{\mathrm{CVaR}_\alpha} = \{f : \Omega \to \mathbb{R} |\ 0 \leq f \leq \frac{1}{\alpha}, \mathbb{E}[f] = 1\}$, which is uniformly integrable (and hence $\sigma(L^1, L^\infty)$-compact). It is shown in Delbaen [?] that a law invariant coherent (convex in general) risk measure on L^∞ is continuous from below if and only if its extension to L^1 takes finite value for some position which is unbounded from below. This is important to know because we will see that any coherent risk measure defined on L^p which can be extended to L^1 is incompatible with pricing rules induced by unbounded stochastic discount factors (like one given by the Black-Scholes model).*

4.3. Compatibility and Good Deals

This section will be devoted to introduce the notion of compatibility between a coherent risk measure and a sub-linear pricing rule and its relation with Good Deals.

Definition 4.3.1. *Let π be a sub-linear pricing rule and ρ a coherent risk measure. We say π and ρ are compatible if there is no sequence $(X_n)_{n=1}^\infty \subset \mathcal{X}$ such that the following conditions simultaneously hold*

$$\pi\left(X_n\right) \leq 0, \ \ \forall n \in \mathbb{N} \tag{4.3.1}$$

$$\lim_{n\to\infty} \rho\left(X_n\right) = -\infty. \tag{4.3.2}$$

We say π and ρ are incompatible if they are not compatible.

As one can see if π and ρ are incompatible, then every manager who uses ρ to assess the risk can make the risk as negative as he/she wishes, which does not *make any economical sense. For further discussion we refer the reader to Balbás et al. [14].*

Now we give our definition of a Good Deal inspired by definitions in Černý and Hodges [19] and Cherny [25].

Definition 4.3.2. *A Good Deal is a position $X \in \mathcal{X}$ such that $\pi(X) \leq 0$ and $\rho(X) < 0$. No Good Deal is an assumption when there is no Good Deal.*

Theorem 4.3.1. *Let ρ be a coherent risk measure and π a sub-linear pricing rule . Let*

$$\mathcal{R} := \left\{ Z \in L^q_+ \middle| \mathbb{E}[Z] = 1,\, \pi\left(X\right) - \mathbb{E}\left(XZ\right) \geq 0, \forall X \in \mathcal{X} \right\}. \tag{4.3.3}$$

The No Good Deal assumption holds if and only if

$$\Delta_\rho \cap \mathcal{R} \neq \emptyset.$$

PROOF. This is easily concluded by using Theorem 3.4 in Cherny [25]. □

Here we try to present an example of a Good Deal, illustrating how these pathological positions could appear in a market.

***Example**: Let Y be a random variable in $L^1 \setminus L^2$. Without loss of generality one can consider that Y is bounded above by a positive number M (otherwise one can pick either $-|Y|1_{\{Y<0\}} + M$ or $-|Y|1_{\{Y\geq 0\}} + M$ in lieu of Y). Let ρ be any law invariant risk measure on L^∞ i.e., for all X, $\rho(X)$ is a function of density of X. Since ρ is law invariant, it is finite on L^1 (see Remark 4.3.1) which implies that $\rho(Y) < \infty$. Let $X_n = Y1_{\{Y\geq -n\}} + \rho(Y)$ and note that $\rho(X_n) \uparrow 0$. Let $\mathcal{X} = L^\infty$ and define $\pi(X) = \mathbb{E}(\frac{1}{\|Y\|_{L^1}}|Y|X)$. Considering the above notations, we have*

$$\begin{aligned}
\pi(X_n) &= \pi(Y1_{\{Y\geq -n\}} + \rho(Y)) \\
&= \pi(Y1_{\{Y\geq -n\}} - M) + M + \rho(Y) \\
&= \pi(-|Y1_{\{Y\geq -n\}} - M|) + M + \rho(Y) \\
&\leq \frac{1}{\|Y\|_{L^1}}\mathbb{E}(-Y^2 1_{\{Y\geq -n\}}) + 2M + \rho(Y) \xrightarrow[n\to\infty]{} -\infty,
\end{aligned}$$

where in the limit we used the fact that $Y \notin L^2$. One can see that for a large enough integer number n_0, there is a position X_{n_0} such that $\rho(X_{n_0}) \leq 0$ whereas

$\pi(X_{n_0}) < 0$. *By definition* X_{n_0} *is a Good Deal. In addition, according to Definition 4.3.1,* ρ *and* π *are incompatible (using* X_n *in the definition).*

4.3.1. A Hedging Problem

Here we consider a more practical discussion when we want to hedge a financial position g *with all possible choices we can make subject to a given budget constraint over a set* $\mathcal{X}$*. This problem will help us to better discover the relation between the concepts of incompatibility and Good Deals.*

Let us consider the following problem

$$\begin{cases} \min \rho\,(X-g)+c \\ \pi\,(X) \leq c \\ X \in \mathcal{X}\,,\, c \in \mathbb{R}. \end{cases} \tag{4.3.4}$$

This problem has been studied in Balbás et al.[12], Balbás et al. [14] and Balbás et al. [15]. The dual of problem (4.3.4) *is found in Balbás et al. [12] as*

$$\begin{cases} \max\ \mathbb{E}[gZ] \\ Z \in \Delta_\rho \cap \mathcal{R} \end{cases} . \tag{4.3.5}$$

Following the discussions in Balbás et al. [12], Balbás et al. [14] and Balbás et al. [15] we have the following theorem

Theorem 4.3.2. *The following statements are equivalent:*

(1) π *and* ρ *are compatible.*

(2) $\mathcal{R} \cap \Delta_\rho \neq \emptyset$.

(3) Problem (4.3.4) *is bounded.*

(4) Problem (4.3.5) *has a feasible solution.*

(5) There is no duality gap between (4.3.4) *and* (4.3.5)*.*

As one can see (4.3.5) *has a solution if and only if* $\Delta_\rho \cap \mathcal{R} \neq \emptyset$*, which obviously reminds us of Theorem 4.3.1. Now we add the following statements to Theorem 4.3.2*

Theorem 4.3.3. *All statements of Theorem 4.3.2 are equivalent to the followings:*

(1) The No Good Deal assumption holds.

(2) $\rho + \pi \geq 0$.

Proof. From Definition 4.3.2, it is obvious that the No Good Deal assumption holds iff for all X in $\mathcal{X}$, $\pi(X) \leq 0$ implies $\rho(X) \geq 0$. Therefore, since $\pi(X - \pi(X)) = 0$, we have that $\rho(X - \pi(X)) \geq 0$. Since ρ is translation invariant, we conclude that $\rho(X) + \pi(X) \geq 0$, showing that 1 implies 2. Now we prove the other implication. To this end, let us suppose that there exists a Good Deal $X \in \mathcal{X}$. By Definition 4.3.2, there exists $X \in \mathcal{X}$ such that $\rho(X) < 0$ and $\pi(X) \leq 0$ which implies $\rho(X) + \pi(X) < 0$. $\square$

In the following remark we show that Good Deals are not rare positions.

Remark 4.3.1. *Suppose that $p \neq 1$.Let ρ be a law invariant coherent risk measure on L^p i.e. $\rho(X) = \rho(Y)$ for any two random variables X, Y with identical distributions. It has recently been proven in Filipovic and Svindland [39] that every law invariant coherent risk measure on L^∞ can canonically be extended to L^1. Let us for a moment denote this extension with $\tilde{\rho}$. According to previous discussions, $\Delta_{\tilde{\rho}}$ is a $\sigma(L^\infty, L^1)$-closed convex set of L^∞. Since $L^p \subseteq L^1$, Δ_ρ is also $\sigma(L^q, L^p)$-closed convex set of L^q. This implies that $\tilde{\rho}$ restricted to L^p can be represented as $\tilde{\rho}(X) = \sup_{\Delta_{Z \in \tilde{\rho}}} \mathbb{E}[-XZ]$ which by $\sigma(L^q, L^p)$-closeness of $\Delta_{\tilde{\rho}}$ implies that $\rho(X) = \sup_{\Delta_{Z \in \Delta_\rho}} \mathbb{E}[-XZ]$. This shows that $\Delta_\rho = \Delta_{\tilde{\rho}}$. Now according to Theorem 4.3.1, this shows that a law invariant risk measure (like CVaR) with a pricing model which has unbounded stochastic discount factors (like the Black-Sccholes model) produces Good Deals.*

4.4. Risk Modification

Discussions in the Remark 4.3.1 show that compatibility may fail in very important cases. This motivates us to modify risk measures to ones compatible with pricing rules.

Definition 4.4.1. *With the same notation as above, let π be a sub-linear pricing rule on $\mathcal{X} \subseteq L^p$, and ρ a coherent risk measure on L^p. A minimal compatible modification, denoted by ρ_m, is a coherent risk measure on L^p such that:*

a) π and ρ_m are compatible, and $\rho \leq \rho_m$;

b) ρ_m is minimal, i.e. for any risk measure $\tilde{\rho}$ such that π and $\tilde{\rho}$ are compatible and $\rho \leq \tilde{\rho} \leq \rho_m$, we have that $\tilde{\rho} = \rho_m$.

Note that the minimal compatible modification is not necessarily unique.

To study the existence of minimal compatible modification we need the following notation. For a given $Z \in L^q \setminus \Delta_\rho$ *let*

$$C(Z) := co(\{Z\} \cup \Delta_\rho). \tag{4.4.1}$$

where co denotes the convex hull. It is easy to see that since Δ_ρ *is* $\sigma(L^q, L^p)$*-compact then* $C(Z)$ *is* $\sigma(L^p, L^q)$*-closed. Define* $\preceq$ *for two members* $Z_1, Z_2 \in L^q \setminus \Delta_\rho$*:*

$$Z_1 \preceq Z_2 \;\Leftrightarrow\; C(Z_1) \subseteq C(Z_2). \tag{4.4.2}$$

Equivalently

$$Z_1 \preceq Z_2 \;\Leftrightarrow\; Z_1 \in C(Z_2). \tag{4.4.3}$$

This relation shows that $\preceq$ *is a transitive relation and then a partial ordering. In the following theorem we see that if* $\mathcal{R} \cap \Delta_\rho = \emptyset$*, the partial ordering* $\preceq$ *has at least one minimal.*

Theorem 4.4.1. *Suppose that* $\mathcal{R} \cap \Delta_\rho = \emptyset$*. Then there exists a minimal point* $Z \in (\mathcal{R}, \preceq)$*.*

Before proving the theorem we need to prove the following lemma

Lemma 4.4.1. *Let* $\{Z_n\}_n$ *be a sequence in* $\mathcal{R}$ *such that* $Z_1 \succeq Z_2 \succeq Z_3 \succeq \ldots$ *and* $Z_n \to Z$ *in* $\sigma(L^q, L^p)$*. Then*

$$\cap_{n\in\mathbb{N}} C(Z_n) = C(Z). \tag{4.4.4}$$

Proof. Fix an arbitrary integer number $N \in \mathbb{N}$. By our assumption we have $Z_n \preceq Z_N$, $\forall n \geq N$ which in turn yields $Z_n \in C(Z_N)$, $\forall n \geq N$. Since $C(Z_N)$ is closed and N is arbitrarily chosen, we deduce that $Z \in C(Z_N)$. That gives for all $N \geq 1$, $C(Z) \subseteq C(Z_N)$ which yields $C(Z) \subseteq \cap_{n\in\mathbb{N}} C(Z_n)$, showing that the right hand-side of (4.4.4) is included in the left hand-side.

In order to prove the other inclusion let $\tilde{Z}$ be a member of $\cap_{n\in\mathbb{N}} C(Z_n)$. For any $n \in \mathbb{N}$, by definition of $C(Z_n)$ there exists $\lambda_n \in [0,1]$ and $Z_n^* \in \Delta_\rho$ such that

$$\tilde{Z} = (1 - \lambda_n) Z_n + \lambda_n Z_n^*.$$

Since Δ_ρ is $\sigma(L^q, L^p)$-compact and $[0,1]$ is bounded, one can extract convergent subsequences from Z_n^* and λ_n converging to $Z^* \in \Delta_\rho$ and $\lambda \in [0,1]$ respectively. In the limit we have

$$\tilde{Z} = (1-\lambda)Z + \lambda Z^*,$$

which means that $\tilde{Z}$ belongs to the convex hull of Z and Δ_ρ. By definition of $C(Z)$ this gives that $\tilde{Z} \in C(Z)$. $\square$

Proof of Theorem 4.4.1 *Fix a member $\bar{Z}$ of $\mathcal{R}$ and let*

$$\mathcal{A} = \left\{ Z \in C(\bar{Z}) \cap \mathcal{R} \,\middle|\, Z \preceq \bar{Z} \right\}. \tag{4.4.5}$$

We show that $\big(\mathcal{A}, \succeq\big)$ satisfies the conditions of Zorn's Lemma. Since $\bar{Z} \in C(\bar{Z})$, the set $\mathcal{A}$ is obviously nonempty. On the other hand let $\{Z_n\}_n$ be a chain in $\mathcal{A}$ i.e. $Z_1 \succeq Z_2 \succeq \dots$. Since $\mathcal{A}$ is $\sigma(L^q, L^p)$-compact, there exists a subsequence $\{Z_{n_k}\}_k$ such that $Z_{n_k} \to Z$ in $\sigma(L^q, L^p)$, for some $Z \in \mathcal{A}$. By applying Lemma 4.4.1 and using the fact that $C(Z_1) \supseteq C(Z_2) \supseteq \dots$ we have that $\cap_{i\in\mathbb{N}} C(Z_i) = C(Z)$. This means that Z is a supremal point of the chain. By applying Zorn's Lemma, there exists a $\preceq$-minimal point $Z \in \mathcal{A}$.

Now we claim that Z is a minimal point for $\mathcal{R}$. Suppose there exists $\tilde{Z}$ in $\mathcal{R}$ such that $\tilde{Z} \preceq Z$. Since $\tilde{Z} \preceq Z \preceq \bar{Z}$, and since $\preceq$ is transitive (see (4.4.3)*) we have that $\tilde{Z} \in C(\bar{Z})$ which by definition gives $\tilde{Z} \in \mathcal{A}$. Since Z is a minimal point for $(\mathcal{A}, \preceq)$ consequently $Z = \tilde{Z}$ which implies that Z is minimal for $(\mathcal{R}, \preceq)$.*

Now the proof of the following theorem is straightforward

Theorem 4.4.2. *Suppose that the No Good Deal assumption does not hold. The risk measure ρ_m is a minimal compatible modification of ρ if and only if*

$$\Delta_{\rho_m} = C(Z)$$

for some minimal Z in $\big(\mathcal{R}, \preceq\big)$.

The following corollary gives a perfect geometrical description of a minimal compatible extension of a coherent risk measure modifying the results in Balbás and Balbás [13].

By Theorems 4.4.1 and 4.4.2 we have the following corollary:

Corollary 4.4.1 (Minimal Modification)**.** *Suppose that the No Good Deal assumption does not hold and ρ_m is a minimal modification of ρ. Then*

$$\rho_m(X) = \max\left\{\rho(X), -\mathbb{E}(ZX)\right\}$$

for some minimal point Z in $(\mathcal{R}, \preceq)$.

4.5. Modification Rules

In the following discussions we propose two major methods for finding a minimal compatible modification ρ_m of ρ. The first method relies on minimizing a third function ϕ, which is interpreted as a spread criteria. This new measure ϕ concerns the fundamentals of the ρ-user. For instance, we will see, by considering $\phi(.) = \|.\|_{L^1}$, that Δ_ρ does not spread out very far in terms of the L^1-norm.

As for the second proposed way of modifying the risk measure, our method is an outcome of finding the No Better Choice (NBC) pricing rule of the Global/Local Efficiency Ratio (see Cherny [25]). A Global/Local Efficiency Ratio is a performance ratio which takes the market fundamentals as well as the risk user desires into account.

4.5.1. Minimal Risk Spread

Let us start with the following definition

Definition 4.5.1. *A function $\phi : L^q \to \mathbb{R}$ is a spread criteria if*

(ϕ1) ϕ is positive and convex.

(ϕ2) The function $(Z, Z_1) \mapsto \phi(Z - Z_1)$ attains its minimum at a point $(Z_{min}, Z^*) \in \mathcal{R} \times \Delta_\rho$.

(ϕ3) The equality $\phi(Z) = 0$ holds if and only if $Z = 0$.

The following theorem enables us to find a minimal compatible modification of a coherent risk measure ρ based on a spread criteria ϕ.

Theorem 4.5.1. *Suppose that the No Good Deal assumption does not hold. Then, in the above notation Z_{min} is a minimal point for $(\mathcal{R}, \preceq)$.*

Proof. Since No Good Deal assumption does not hold, by Theorem 4.3.3 we know that $\mathcal{R} \cap \Delta_\rho = \emptyset$. To prove the theorem's statement we suppose, to the contrary, that Z_{min} is not minimal. Then there exists $\tilde{Z} \in \mathcal{R}$ such that $\tilde{Z} \in C(Z_{min})$ and $\tilde{Z} \neq Z_{min}$. Since $Z_{min} \neq \tilde{Z} \in C(Z_{min})$, by definition there exists $\lambda \in (0,1]$ and $Z_1 \in \Delta_\rho$ such that

$$\tilde{Z} = (1-\lambda)Z_{min} + \lambda Z_1.$$

By convexity of Δ_ρ we know that $Z_2 = (1-\lambda)Z^* + \lambda Z_1 \in \Delta_\rho$. Given assumptions $(\phi 1), (\phi 3)$ we have

$$\begin{aligned}\phi(\tilde{Z} - Z_2) &= \phi\Big((1-\lambda)Z_{min} + \lambda Z_1 - ((1-\lambda)Z^* + \lambda Z_1)\Big)\\ &= \phi\Big((1-\lambda)(Z_{min} - Z^*)\Big)\\ &\leq (1-\lambda)\phi(Z_{min} - Z^*).\end{aligned}$$

Since $0 \leq 1-\lambda < 1$, by definition of Z_{min} we have that $\phi(Z_{min} - Z^*) = 0$. By condition $(\phi 3)$ we get that $Z_{min} = Z^*$ which contradicts our Good Deal assumption. □

4.5.1.1. *Compatible Conditional Value at Risk (CCVaR)*

In this part we are going to use the theory we have developed in the last section by implementing $\phi(X) = \int_\Omega |X| d\mathbb{P}$ *and* $\rho = \mathrm{CVaR}_\alpha$, *for some confidence level* $\alpha \in (0,1)$ *in Theorem 4.5.1. Interestingly, we will see that in order to find the Compatible CVaR (i.e., compatible with a given* π*), we will have to find a stochastic discount factor with the least European call option price with strike price* $\frac{1}{\alpha}$. *We start with the following lemma.*

Lemma 4.5.1. *For a given* $g \in L^1_+$ *with* $\mathbb{E}[g] = 1$, *the* L^1*-distance between* g *and* $\Delta_{\mathrm{CVaR}_\alpha}$ *equals* $2\int_\Omega \left(g - \frac{1}{\alpha}\right)^+$ *i.e.*

$$\min_{Z \in \Delta_{\mathrm{CVaR}_\alpha}} \int_\Omega |g - Z| = 2\int_\Omega \left(g - \frac{1}{\alpha}\right)^+.$$

Furthermore, the minimum is attained only in points Z^* *given as*

$$Z^* = \frac{1}{\alpha} 1_{\{g \geq \frac{1}{\alpha}\}} + (g+h)1_{\{g < \frac{1}{\alpha}\}}, \tag{4.5.1}$$

where h *is a non-negative function for which* $(g+h)1_{\{g<\frac{1}{\alpha}\}} \leq \frac{1}{\alpha}$ *and*

$$\int_{\{g<\frac{1}{\alpha}\}} h = \int_{\Omega} \left(g - \frac{1}{\alpha}\right)^{+}. \tag{4.5.2}$$

Proof. First recall from Rockafellar et al. [58] that

$$\Delta_{\mathrm{CVaR}_\alpha} = \left\{ f \in L^1 \middle| 0 \le f \le \frac{1}{\alpha}, \mathbb{E}[f] = 1 \right\}.$$

Let $Z \in \Delta_{\mathrm{CVaR}_\alpha}$ and define

$$\begin{aligned}
Z_1 &:= (Z - g)1_{Z \ge g}, \\
Z_2 &:= (g - Z)1_{\{g \ge Z, g < \frac{1}{\alpha}\}}, \\
Z_3 &:= \min(Z, g), \\
Z_4 &:= (\frac{1}{\alpha} - Z)1_{g \ge \frac{1}{\alpha}}, \\
Z_5 &:= (g - \frac{1}{\alpha})1_{g \ge \frac{1}{\alpha}}.
\end{aligned}$$

It is clear that

$$\begin{aligned}
&Z_1 + Z_3 = Z, \\
&g = Z_2 + Z_3 + Z_4 + Z_5.
\end{aligned}$$

Therefore,

$$\int Z_1 + \int Z_3 = 1, \tag{4.5.3}$$

$$1 = \int Z_2 + \int Z_3 + \int Z_4 + \int Z_5. \tag{4.5.4}$$

On the other hand, since Z_2 and Z_4 are non-negative we have

$$2\int Z_2 + \int Z_4 \ge 0. \tag{4.5.5}$$

Combining (4.5.5) with (4.5.3) and (4.5.4), yields

$$\int Z_1 + \int Z_2 \ge \int Z_5.$$

Adding one more $\int Z_5$ to both sides of the last inequality, we obtain

$$\int Z_1 + \int Z_2 + \int Z_5 \ge 2\int Z_5.$$

Having this, one can see that

$$\int |Z-g| = \int Z_1 + \int Z_2 + \int_{g\geq\frac{1}{\alpha}} (g-Z) \tag{4.5.6}$$

$$\geq \int Z_1 + \int Z_2 + \int Z_5 \tag{4.5.7}$$

$$\geq 2\int Z_5 = 2\int \left(g-\frac{1}{\alpha}\right)^+. \tag{4.5.8}$$

Therefore, $2\int \left(g-\frac{1}{\alpha}\right)^+$ is smaller than $\int |Z-g|$ for all Z.

Now we take three steps to conclude the proof: First, we show that at least one Z^* exists. Second, we show that every Z^* introduced in (4.5.1) is a minimal point. Third, we prove that every minimal point has the same structure as in (4.5.1).

Step 1. We show that there exists a function h which satisfies the conditions in Lemma 4.5.1 and can be put into (4.5.1).

Observe that since $0 > 1-\frac{1}{\alpha} = \int (g-\frac{1}{\alpha}) = \int_{g\geq\frac{1}{\alpha}}(g-\frac{1}{\alpha}) + \int_{g<\frac{1}{\alpha}}(g-\frac{1}{\alpha})$ we have $\int_{g\geq\frac{1}{\alpha}}(g-\frac{1}{\alpha}) < \int_{g<\frac{1}{\alpha}}(\frac{1}{\alpha}-g)$. Let $\lambda := \frac{\int_{g\geq\frac{1}{\alpha}}(g-\frac{1}{\alpha})}{\int_{g<\frac{1}{\alpha}}(\frac{1}{\alpha}-g)}$,and note that $\lambda < 1$. Defining $h := \lambda(\frac{1}{\alpha}-g)1_{g<\frac{1}{\alpha}}$, it is clear that h fulfills the conditions of Lemma 4.5.1.

Step 2. Suppose that h is a non-negative function for which $(g+h)1_{\{g<\frac{1}{\alpha}\}} \leq \frac{1}{\alpha}$ and (4.5.2) holds. Define

$$Z^* := \frac{1}{\alpha}1_{g\geq\frac{1}{\alpha}} + (g+h)1_{g<\frac{1}{\alpha}}.$$

First we show that $Z^* \in \Delta_{\mathrm{CVaR}_\alpha}$. By construction it is clear that $0 \leq Z^* \leq \frac{1}{\alpha}$. On the other hand by (4.5.2) we have that

$$\begin{aligned}
\int_\Omega Z^* &= \int_{g\geq\frac{1}{\alpha}} \frac{1}{\alpha} + \int_{g<\frac{1}{\alpha}} g + \int_{g<\frac{1}{\alpha}} h \\
&= \int_{g\geq\frac{1}{\alpha}} \frac{1}{\alpha} + \int_{g<\frac{1}{\alpha}} g + \int_\Omega \left(g-\frac{1}{\alpha}\right)^+ \\
&= \int_{g\geq\frac{1}{\alpha}} \frac{1}{\alpha} + \int_{g<\frac{1}{\alpha}} g + \int_{g\geq\frac{1}{\alpha}} \left(g-\frac{1}{\alpha}\right) \\
&= \int_\Omega g = 1.
\end{aligned}$$

Now we show that $Z_2^* = Z_4^* = 0$. It is easy to see that $Z_4^* = 0$. As for $Z_2^* = 0$, just observe that by definition of Z^*, $\{g < \frac{1}{\alpha}, g \geq Z^*\} = \{h = 0\}$, and therefore

$$Z_2^* = (g - Z^*)1_{\{g \geq Z^*, g < \frac{1}{\alpha}\}} = -h1_{\{g \geq Z^*, g < \frac{1}{\alpha}\}} = -h_{\{h=0\}} = 0.$$

On the other hand, it is also clear that $(g - Z^*)1_{\{g \geq \frac{1}{\alpha}\}} = 0$. Given this, since $Z_2^* = Z_4^* = 0$, we have equalities in (4.5.7) and (4.5.8), which implies that Z^* is a minimal point.

Step 3. Let us denote a minimal point by Z^*. From Steps 1,2 it is clear that the minimum is

$$2\int \left(g - \frac{1}{\alpha}\right)^+.$$

This, along with (4.5.7) and (4.5.8), shows that for any minimal point $Z^* \in \Delta_\rho$ we must have $Z_2^* = Z_4^* = 0$. The equality $Z_4^* = 0$ implies that

$$Z^* 1_{g \geq \frac{1}{\alpha}} = \frac{1}{\alpha}. \tag{4.5.9}$$

This is the first part of (4.5.1).

Let $h := (Z^* - g)1_{\{g < \frac{1}{\alpha}\}}$. By construction, $Z^* 1_{\{g < \frac{1}{\alpha}\}} = (h + g)1_{\{g < \frac{1}{\alpha}\}}$, which is the second part of (4.5.1).

Now we must show that h is non-negative, $(g + h)1_{\{g < \frac{1}{\alpha}\}} \leq \frac{1}{\alpha}$ and that (4.5.2) holds. From $0 = Z_4^* = (g - Z^*)1_{\{g \geq Z^*, g < \frac{1}{\alpha}\}}$ it turns out that g cannot be larger than Z^* on $\{g < \frac{1}{\alpha}\}$. This gives that the function $h = (Z^* - g)1_{\{g < \frac{1}{\alpha}\}}$ is non-negative. Since $Z^* \leq \frac{1}{\alpha}$, it is also clear that $(g + h)1_{\{g < \frac{1}{\alpha}\}} \leq \frac{1}{\alpha}$.

Now by assumption that Z^* is minimal, definition of h and (4.5.9) we have that

$$\begin{aligned} 2\int \left(g - \frac{1}{\alpha}\right)^+ &= \int |g - Z^*| \\ &= \int_{g \geq \frac{1}{\alpha}} \left(g - \frac{1}{\alpha}\right) + \int_{g < \frac{1}{\alpha}} h \\ &= \int \left(g - \frac{1}{\alpha}\right)^+ + \int_{g < \frac{1}{\alpha}} h, \end{aligned}$$

which shows that (4.5.2) hold and the proof is complete.

From Theorem 4.5.1 and Lemma 4.5.1 we deduce the following theorem:

Theorem 4.5.2. *Let SDF be the set of all Stochastic Discount Factors (e.g. EMM in an incomplete market). Suppose that the minimum of* $2\mathbb{E}[(\cdot - \frac{1}{\alpha})^+]$ *over SDF is attained at* $g^* \in$*SDF. Then* g^* *is a minimal point of* $(\mathrm{SDF}, \preceq)$.

Remark 4.5.1. *Interestingly one can see that finding the minimal extension for CVaR is equivalent to finding a stochastic discount factor with the least European call option price with strike* $\frac{1}{\alpha}$.

Remark 4.5.2. *In an incomplete market, there is more than one equivalent martingale measure. Among many choices, the right pick is always an important question. For instance, the minimal martingale measure provided by the Föllmer-Schweizer decomposition, the one which is the nearest in* L^q*-norm to the historical measure* $\mathbb{P}$*, or the one which has the least entropy could be named among many (see Chan [20]). Here, we can add another to this list, which concerns the existence of Good Deals.*

4.5.2. Global Risk and Performance Maximization

In this section we propose the second way of modifying a risk measure which will be carried out via studying the following coherent risk measure:

$$X \mapsto \max\{\rho(X), \pi(-X)\}.$$

We call this risk measure as Global Risk measure and denote by $GR(X)$. Indeed the Global Risk does not only assess the trader's risk, but also the market response to going short on X, which could be interpreted as the market risk. As usual in the literature of coherent risk measure, in the sequel, we will denote the function $-\rho$ by u, and we will call it the *monetary utility* associated with ρ.

For our discussions in this section we need the following assumption on $\mathcal{R}$:

$$\mathcal{R} \text{ is } \sigma(L^q, L^p) - \text{compact.} \tag{4.5.10}$$

Now we start to study the efficiency ratio $\frac{u(X)}{GR(X)}$ in order to propose another way of finding a minimal compatible modification of risk measure ρ. We have the following definition

Definition 4.5.2. *For a couple (π, ρ) the Global/Local performance ratio GL is defined as follows:*

$$GL(X) = \begin{cases} +\infty & \text{if } GR(X) < 0, \\ \frac{u(X)}{GR(X)} & \text{if } GR(X) \geq 0 \text{ and } u(X) > 0, \\ 0 & \text{if } GR(X) \geq 0 \text{ and } u(X) \leq 0, \end{cases} \tag{4.5.11}$$

when $\frac{\text{positive}}{0} = +\infty$.

It is easy to show that

$$GL(X) = \begin{cases} +\infty & \text{if } u(X) > 0 \text{ and } \pi(-X) \leq 0, \\ \frac{u(X)}{\pi(-X)} & \text{if } u(X) > 0 \text{ and } \pi(-X) > 0, \\ 0 & \text{if } u(X) \leq 0. \end{cases} \tag{4.5.12}$$

This is a measure to see how much it is worth to keep X. Further interpretation is left to the reader.

Now let us suppose that the No Good Deal assumption holds. Let X be a financial position such that $\pi(X) \leq 0$. It is clear since $\mathcal{R} \cap \Delta_\rho \neq \emptyset$ then $u(X) \leq 0$, and by (4.5.12) we have $GL(X) = 0$. However, in the opposite case, when the No Good Deal assumption does not hold, i.e. $\mathcal{R} \cap \Delta_\rho = \emptyset$, we always have $\sup_{\pi(X) \leq 0} GL(X) > 0$. This number shows how far a market is from the No Good Deal assumption. This can be summarized in the following proposition

Proposition 4.5.1. *The No Good Deal assumption holds if and only if* $GL(X) = 0$ *for all* X *in* $\{\pi \leq 0\}$.

Here we lead the discussion to the No Better Choice pricing rule associated with the performance ratio GL defined by Cherny [25].

Definition 4.5.3. *For any financial position g the NBC price of g is a real number x such that*

$$\sup_{\{X+h(g-x) \,|\, \pi(X) \leq 0\,,\, h \in \mathbb{R}\}} GL\Big(X + h(g-x)\Big) = \sup_{\{X \,|\, \pi(X) \leq 0\}} GL(X). \tag{4.5.13}$$

Actually it is the cost for g in which the maximum efficiency ratio does not increase by adding the new product g. The set of all NBC prices are denoted by I_{NBC}.

We denote the supremum in (4.5.13) by R^*, i.e.

$$R^* = \sup_{\{X \mid \pi(X) \le 0\}} GL(X).$$

In Cherny [25] it is shown that

$$R^* = \inf\left\{ R \ge 0 \,\middle|\, \left(\frac{1}{1+R}\Delta_\rho + \frac{R}{1+R}\bar{co}\big(\Delta_\rho \cup \mathcal{R}\big) \right) \cap \mathcal{R} \ne \emptyset \right\}.$$

Since Δ_ρ and $\mathcal{R}$ are $\sigma(L^q, L^p)$-compact and both Δ_ρ and $\mathcal{R}$ are convex we get

$$\bar{co}\big(\Delta_\rho \cup \mathcal{R}\big) = co\big(\Delta_\rho \cup \mathcal{R}\big). \tag{4.5.14}$$

To show (4.5.14) let $X \in co\big(\Delta_\rho \cup \mathcal{R}\big)$. Then, $X = \sum_{i=1}^{k} \mu_i Y_i + \sum_{j=1}^{l} \lambda_j Z_j$, where the μ_i, λ_j are positive with $\sum \mu_i + \sum \lambda_j = 1$, and also $(Y_i, Z_j) \in \Delta_\rho \times \mathcal{R}$ for $1 \le i \le k$, $1 \le j \le l$. Letting $\mu = \sum \mu_i$ and $\lambda = \sum \lambda_j$ we have that

$$X = \mu\bigg(\sum \frac{\mu_i}{\mu} Y_i \bigg) + \lambda\bigg(\sum \frac{\lambda_j}{\lambda} Z_j \bigg).$$

By convexity of Δ_ρ and $\mathcal{R}$ it follows that every member of $co\big(\Delta_\rho \cup \mathcal{R}\big)$ can be written as $X = \mu Y + \lambda Z$ for $(Y, Z) \in \Delta_\rho \times \mathcal{R}$ where $\lambda + \mu = 1$, $\mu, \lambda \ge 0$. Now let us suppose that $X_n \in co\big(\Delta_\rho \cup \mathcal{R}\big)$ converges in $\sigma(L^q, L^p)$ to X. There exist $0 \le \lambda_n \le 1$, $Y_n \in \Delta_\rho$ and $Z_n \in \mathcal{R}$ such that $X_n = (1 - \lambda_n)Y_n + \lambda_n Z_n$. Since Δ_ρ and $\mathcal{R}$ are $\sigma(L^q, L^p)$-compact, upon taking subsequences one can assume that Y_n, Z_n and λ_n converge to Y, Z and λ respectively in $\Delta_\rho, \mathcal{R}$ and $[0, 1]$. This implies that $X = (1 - \lambda)Y + \lambda Z \in co\big(\Delta_\rho \cup \mathcal{R}\big)$.

By (4.5.14) and $\Delta_\rho \cap \mathcal{R} = \emptyset$ it is clear that the expression $\frac{1}{1+R} Z_1 + \frac{R}{1+R} Z \in \mathcal{R}$, for some positive number $R > 0$ and for some $Z_1 \in \Delta_\rho$ and $Z \in co(\Delta_\rho \cup \mathcal{R})$, implies that $Z \in \mathcal{R}$. This implies that R^* can be rewritten as follows

$$R^* = \inf\left\{ R \ge 0 \,\middle|\, \left(\frac{1}{1+R}\Delta_\rho + \frac{R}{1+R}\mathcal{R} \right) \cap \mathcal{R} \ne \emptyset \right\}.$$

Let

$$\mathcal{D}^* = \frac{1}{1+R^*}\Delta_\rho + \frac{R^*}{1+R^*}\bar{co}\big(\Delta_\rho \cup \mathcal{R}\big).$$

In Cherny [25] it is shown that

$$I_{NBC}(g) = \big\{ \mathbb{E}(Zg) \big|\, Z \in \mathcal{D}^* \big\}.$$

Let us associate with each Z the following number

$$r_Z = \inf\left\{R \geq 0 \,\middle|\, \exists(Z_1, \tilde{Z}) \in \Delta_\rho \times co(\Delta_\rho \cup \mathcal{R}), \frac{1}{1+R}Z_1 + \frac{R}{1+R}\tilde{Z} = Z\right\} \tag{4.5.15}$$

As discussed in Corollary 3.10 Cherny [25], $\mathcal{D}^* \cap \mathcal{R}$ consists of all points in $\mathcal{R}$ with minimum r_Z. By discussion above it is now clear that (4.5.15) equals

$$\inf\left\{R \geq 0 \,\middle|\, \exists(Z_1, \tilde{Z}) \in \Delta_\rho \times \mathcal{R}, \frac{1}{1+R}Z_1 + \frac{R}{1+R}\tilde{Z} = Z\right\}$$

Let

$$\begin{cases} d : \Delta_\rho \times \mathcal{R} \to [0, +\infty], \\ d(Z_1, Z) = \inf\left\{R \geq 0 \,\middle|\, \exists\tilde{Z} \in \mathcal{R},\ \frac{1}{1+R}Z_1 + \frac{R}{1+R}\tilde{Z} = Z\right\}. \end{cases}$$

To see $d(Z_1, Z)$ geometrically, we connect Z_1 to Z and continue until hitting the last point in $\mathcal{R}$, named $\tilde{Z}$ (since $\mathcal{R}$ is $\sigma(L^q, L^p)$-compact, the last point exists). So there exists $R \geq 0$ such that $Z = \frac{1}{1+R}Z_1 + \frac{R}{1+R}\tilde{Z}$. Then $d(Z_1, Z) = R$. In the case that the continuation of the semi line $\overrightarrow{Z_1 Z}$ hits $\mathcal{R}$ only in Z (i.e. $Z = \tilde{Z}$) we put $d(Z_1, Z) = +\infty$. The function d is lower semi-continuous.

Lemma 4.5.2. *The function d defined above is $\sigma(L^q, L^p)$-lower semi-continuous.*

Proof. To show that d is $\sigma(L^q, L^p)$-lower semi-continuous we have to prove that

$$C_a = \left\{(Z_1, Z) \in \Delta_\rho \times \mathcal{R} \,\middle|\, d(Z_1, Z) \leq a\right\}$$

is $\sigma(L^q, L^p)$-closed for every positive number $a \in [0, \infty]$. To this end let fix $a \in [0, \infty]$ and let $\{(Z_1^n, Z^n)\}_n$ be a sequence in C_a, converging to $(Z_1, Z) \in \Delta_\rho \times \mathcal{R}$ in $\sigma(L^q, L^p)$. The case $a = +\infty$ is trivial. The case $a = 0$ is never applied since we are assuming that $\Delta_\rho \cap \mathcal{R} = \emptyset$. So let us suppose that $a \in (0, +\infty)$. For each n there exists $\tilde{Z}^n$ such that $Z^n = \frac{1}{1+d(Z_1^n, Z^n)}z_1^n + \frac{d(Z_1^n, Z^n)}{1+d(Z_1^n, Z^n)}\tilde{Z}^n$. Since $d(Z_1^n, Z^n)$ is bounded, by $\sigma(L^q, L^p)$-compactness of $\mathcal{R}$ one can find subsequence n_k such that $d(Z_1^{n_k}, Z^{n_k})$ and $\tilde{Z}^{n_k}$ converge respectively to d $(0 \leq d \leq a)$ and $\tilde{Z} \in \mathcal{R}$.

In the limit we have that $\frac{1}{1+d}Z_1 + \frac{d}{1+d}\tilde{Z} = Z$, which by definition in turn yields $d(Z_1, Z) \leq d \leq a$.

□

As mentioned above, by Corollary 3.10 Cherny [25] and Lemma 4.5.2 one can deduce that

$$\mathcal{D}^* \cap \mathcal{R} = \left\{ Z \in \mathcal{R} \,|\, \exists Z_1 \in \Delta_\rho \,,\, d(Z_1, Z) \text{ is minimal } \right\}.$$

The members of the set $\mathcal{D}^* \cap \mathcal{R}$ are the discount factors for the No Better Choice pricing technique. But interestingly the members of this set are also minimal for $(\mathcal{R}, \preceq)$ (see the next theorem) which by Theorem 4.4.2 leads us to a good choice of the risk recovery.

Theorem 4.5.3. *All members of $\mathcal{D}^* \cap \mathcal{R}$ are minimal for $(\mathcal{R}, \preceq)$.*

Proof Let $Z \in \mathcal{D}^* \cap \mathcal{R}$. By Lemma 4.5.2 we can suppose that there exists $Z_1^{min} \in \Delta_\rho$ such that $d(Z_1^{min}, Z)$ is minimal over $\Delta_\rho \times \mathcal{R}$. From discussions above we know that there exists $\tilde{\tilde{Z}} \in \mathcal{R}$ such that

$$Z = \frac{1}{1 + d(Z_1^{min}, Z)} Z_1^{min} + \frac{d(Z_1^{min}, Z)}{1 + d(Z_1^{min}, Z)} \tilde{\tilde{Z}}.$$

Now let us suppose, to the contrary, that there exists $\tilde{Z} \in C(Z) \cap \mathcal{R}$ and $\tilde{Z} \neq Z$. By definition there exists $Z_2 \in \Delta_\rho$ and $R \in [0, +\infty)$ such that $\frac{1}{1+R}Z_2 + \frac{R}{1+R}Z = \tilde{Z}$. From this relation it turns out that $d(Z_2, \tilde{Z}) \leq R$ which yields $d(Z_1^{min}, Z) \leq R < +\infty$. This assures us that $Z \neq \tilde{\tilde{Z}}$.

Since Z is convex combination of Z_1^{min} and $\tilde{\tilde{Z}}$ the three points $Z_1^{min}, Z, \tilde{\tilde{Z}}$ are on the same direction. We claim that the point $\tilde{Z}$ cannot be on the line that passes through $Z_1^{min}, Z, \tilde{\tilde{Z}}$. In order to see this, first note that since $\tilde{Z} \prec Z \prec \tilde{\tilde{Z}}$ we have that $\tilde{Z} \notin \overrightarrow{Z\tilde{\tilde{Z}}}$. Hence, if $\tilde{Z}$ lies on the same direction as $Z_1^{min}, Z, \tilde{\tilde{Z}}$ two possibilities exist: either $\tilde{Z} \in [Z_1^{min}, Z)$ or $Z_1^{min} \in [\tilde{Z}, Z)$. The first is ruled out since obviously in that case $d(Z_1^{min}, \tilde{Z}) < d(Z_1^{min}, Z)$. The second possibility is also ruled out since in that case by convexity of $\mathcal{R}$, we get $Z_1^{min} \in \mathcal{R}$.

Now we have four different points $Z_1^{min}, Z, \tilde{\tilde{Z}}, \tilde{Z}$ which are not in the same direction while three of them, $Z_1^{min}, Z, \tilde{\tilde{Z}}$ are. As a result the convex combination of these four points lie in a two dimensional affine space P. It is clear that Z_2

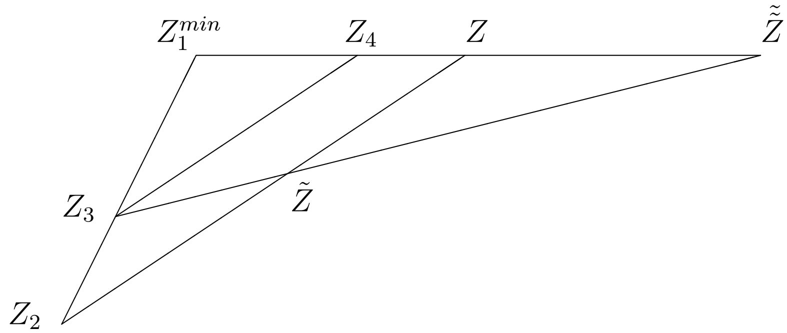

FIGURE 4.1. The proof illustration of Theorem 4.5.3

also belongs to P. Note that $Z_2 \neq Z_1^{min}$, since otherwise $\tilde{Z}$ is on the the line passing through $Z_1^{min}, Z, \tilde{\tilde{Z}}$. In the affine space P, the side ZZ_2 of the triangle $\triangle Z_1^{min} Z Z_2$ is hit by the semi-line $\overrightarrow{\tilde{\tilde{Z}}\tilde{Z}}$ in point $\tilde{Z}$. Therefore, the continuation of $\overrightarrow{\tilde{\tilde{Z}}\tilde{Z}}$ should hit the other side, $Z_1^{min}Z_2$ in a point denoted by Z_3 (the opposite side is impossible since again it puts $\tilde{Z}$ on the line passing through $Z_1^{min}, Z, \tilde{\tilde{Z}}$). By convexity of Δ_ρ, Z_3 belongs to Δ_ρ. Now on the side $Z_1^{min}Z$ of the triangle $\triangle Z_1^{min} Z Z_2$ we find a point Z_4 such that Z_3Z_4 is parallel to Z_2Z. Obviously $Z_4 \in (Z_1^{min}, Z)$. Since Z_3Z_4 and Z_2Z are parallel we have:

$$\frac{|Z_3\tilde{Z}|}{|\tilde{Z}\tilde{\tilde{Z}}|} = \frac{|Z_4 Z|}{|Z\tilde{\tilde{Z}}|} < \frac{|Z_1^{min} Z|}{|Z\tilde{\tilde{Z}}|} = d(Z_1^{min}, Z). \tag{4.5.16}$$

But by definition $d(Z_3, \tilde{Z}) \leq \frac{|Z_3\tilde{Z}|}{|\tilde{Z}\tilde{\tilde{Z}}|}$. Therefore, $d(Z_3, \tilde{Z}) < d(Z_1^{min}, Z)$, which is a contradiction. $\square$

Chapter 5

FURTHER DISCUSSIONS ON GOOD DEALS

Résumé

Dans le chapitre précédent, nous avons examiné en détails le concept de "bonnes affaires". Nous avons vu, en particulier, comment une sous-estimation du risque (ou une sous-capitalisation en contexte financier), pouvait engendrer de "bonnes affaires". Nous avons aussi vu comment l'existence de bonnes affaires pouvait être expliqué en terme d'incompatibilité entre une mesure de risque et d'une règle de tarification. Nous avons discuté de l'élimination de bonnes affaires en ajustant cette incompatibilité (du point de vue de la solvabilité).

L'objectif de ce dernier chapitre est d'attirer l'attention du lecteur sur quelques problèmes rencontrés dans les chapitres antérieurs. Comme nous l'avons vu dans le chapitre 4, section 4.3.1, en travaillant avec une mesure de risque cohérente et invariante par rapport à la loi de la variable, nous nous retrouvons toujours, pour certains modèles, avec de bonnes affaires. Dans ce chapitre, nous allons dans un premier temps étendre la notion de bonnes affaires à une plus grande famille de mesures de risque et de règles de tarification (section 5.2), et deuxièment, nous allons tenter de démontrer comment une information imparfaite engendre de bonnes affaires (section 5.3.2). Finalement, nous discuterons comment, dans un marché parfait, le choix de la mesure de risque peut engendrer de bonnes affaires, peu importe la règle de tarification (section 5.3.3). Cependant, avant de commencer, nous allons montrer coment de bonnes affaires peuvent être observées en pratique.

Astract

In the previous chapter we looked into the concept of a Good Deal in some detail. Indeed, we discussed how underestimating the risk (or in financial terminology under-capitalization) could produce a Good Deal. We also showed how existence of Good Deals can be explained in terms of the incompatibility between a risk measure and a pricing rule. We also discussed how we can deal with incompatibility in such a way that Good Deals are ruled out. We looked at incompatibility from a solvency perspective.

The objective of this final chapter is to bring to the reader's attention several issues that arise naturally after previous discussions. As we have seen in Chapter 4 Section 4.3.1, dealing with a law invariant coherent risk measure, in some well-known models, we always end up having Good Deals. In this chapter we will first, extend the concept of a Good Deal for a larger family of risk measures and pricing rules (5.2) and second, try to show how imperfect information produces a Good Deal (Section 5.3.2). Finally we discuss how in a perfect market, the choice of a risk measure can produce Good Deals, regardless of the choice of the pricing rule (Section 5.3.3). However, before that we show how a Good Deal can be observed in a real life practice.

5.1. An Example

In this section we give an example from the market which we believe can be interpreted as a Good Deal.

Our example is a particular financial product, a Credit Default Swap (hence CDS), which is commonly used for hedging against default of a bond. After the years 2001 and 2002, according to some decisions made by the Federal Reserve and the Treasury in the US, the housing market (or real estate market) started to grow. That motivated financial institutions to issue bonds backed by mortgages. Very soon the bonds backed by mortgages became very popular among financial practitioners. The popularity of those bonds was due to a belief that the prices in *the housing market almost never drop. Of course, this belief was supported with available data from the past. Accordingly, many derivatives started to be issued on those bonds, among which one can name CDS. A CDS is an insurance contract on a bond in which the insurer accepts to compensate the loss of defaulting a bond subject to receiving a regular payment, up to either default or the end of the contract. Insurance companies (like AIG) started to issue those kind of insurance contracts and sell those to customers. Those products seemed to be very good deals for insurance companies because according to previous data, the bonds backed by mortgages had almost never defaulted (due to the increasing trend of house prices), while they also payed back regular payments.*

In mathematical terms, let T *be the time that a CDS contract ends and let* θ *be the time the bond defaults. An insurance company is payed for sure by* $p \times (\theta \wedge T)$*, where* p *is the amount of the regular payments. Since the insurance company considers that the bond never defaults, i.e.* $\theta = \infty$*, the amount the company will receive is supposed to be* $p \times T > 0$*. This implies that this contract costs* $-p \times T$ *to the insurance company (indeed it pays back so it has negative price). Also, by considering no default, the risk becomes zero* $Risk(CDS) = 0$ *(whatever the risk measure is). Now let us look at* $CDS + p \times T$*. This is a product which has a negative risk,* $Risk(CDS + p \times T) = Risk(CDS) - p \times T = 0 - p \times T < 0$ *(here we consider a translation invariant function as Risk) while also* $CDS + p \times T$ *has zero price,* $Price(CDS + p \times T) = Price(CDS) + p \times T = -p \times T + p \times T = 0$ *(also we consider a translation invariant function as Price). This shows that* $CDS - p \times T$ *is a Good Deal, which is so, of course from the insurance company point of view.*

The CDS we mention above is not an arbitrage. Actually, the fact is there were some moments in US history when the price in the housing market dropped, but those moments are very short and the drops were not very big. What makes a CDS a Good Deal is the assessment of the default-risk based on previous data. Indeed, according to previous data the risk of default of the bonds backed with mortgages is negligible or even zero. The other fact is that the information about the short drops in the housing market was "publicly available" but what may ignore that information is the process of risk assessment which depends on the model and *the risk measure. In other words, not only the information, but also the tools we use in our risk assessment may be a source of producing Good Deals. These are point which will be discussed in upcoming sections, but first we try to extend the definition of Good Deals to a wider family of risk measures which contains a more practical risk measures like Value at Risk.*

5.2. Extending the Notion of Good Deals

In this section we see how one can extend the definition of a Good Deal in several different directions. Each direction is aimed at involving new families of risk measures which previously we could not bring into our discussion, for instance natural risk statistics and expectation bounded risks. The following two extensions are motivated from the equivalent statements given in Theorem 4.3.3.

Definition 5.2.1. *Let $\rho : L^p \to \mathbb{R}$ and $\pi : \mathcal{Y} \subseteq L^p \to \mathbb{R}$ be two translation invariant functions, i.e.,*

$$\rho(X + c) = \rho(X) - c \ , \ \forall X \in L^p, c \in \mathbb{R}, \tag{5.2.1}$$

$$\pi(Y + c) = \pi(Y) + c \ , \ \forall Y \in \mathcal{Y}, c \in \mathbb{R}. \tag{5.2.2}$$

We say that the couple (ρ, π) does not produce a Good Deal on $\mathcal{Y}$ if $\rho(X)+\pi(X) \geq 0 \, , \forall X \in \mathcal{Y}$.

This definition includes at least the following four families:

(1) Convex risk measures.

(2) Natural risk statistics.

(3) Coherent risk contribution. Let X, Y be two random variables and $\rho : L^p \to \mathbb{R}$ a coherent risk measure, and consider the following definition of coherent risk contribution:

$$\rho(X; Y) = \lim_{\epsilon \downarrow 0} \frac{\rho(Y + \epsilon X) - \rho(Y)}{\epsilon}. \tag{5.2.3}$$

This definition appears in the works related to capital allocation (with a slight difference), for instance see [40] and [63]. It is very easy to see from the definition that $\rho(X + c; Y) = \rho(X; Y) - c$.

(4) Distortion risk measure. Distortion risk measures are developed from premium principles studied in [65], and are defined as follows:

A distortion function is a non-decreasing function $g : [0,1] \to [0,1]$ with $g(0) = 0$ and $g(1) = 1$. A distortion risk measure associated with distortion function g is defined as

$$\rho_g(X) = \int_0^\infty g(S(t))dt, \tag{5.2.4}$$

where $S(t)$ is the accumulative distribution function of $-X$. This family of risk measures contains VaR_α (when $g(t) = 1_{[1-\alpha,1]}$) and the family of insurance risk measures.

As one can see this definition enables us to study the concepts of a Good Deal in the framework of Chapter 2 and 3. On the other hand, in Section 5.3.3, we will see that Value at Risk (and even a larger family which contains Value at Risk) always produce Good Deals. Of course, as a Good Deal in that section we mean the one we just defined.

In what follows, we introduce another family of risk measures and pricing rules which contain the expectation bounded risk and the deviation risk measures defined in [58] as particular cases. This extension is also important because we can extend very easily Theorem 4.3.3 to that new framework.

Definition 5.2.2. *Let $\rho : L^p \to \mathbb{R}$ be a sub-additive functions and π a sub-additive, translation invariant and positive homogeneous function. Then X is a Good Deal if $\rho(X) < 0$ and $\pi(X) \le 0$ simultaneously hold.*

If ρ and π are lower semi-continuous, they can be represented as follows:

$$\rho(X) = \sup_{Y \in \Delta_\rho} \mathrm{E}[-XY],$$

$$\pi(X) = \sup_{Y \in \mathcal{R}} \mathrm{E}[XY],$$

for $\Delta_\rho \subseteq L^q$, $\mathcal{R} \subseteq L^q$ and $\mathrm{E}[Y] = 1\,,\ \forall Y \in \mathcal{R}$.

This family of risk measures contains the expectation bounded risk measures where $\mathrm{E}[Y] = 1$ for all $Y \in \Delta_\rho$. Following the same idea of proof as in Theorem 4.3.3, one can show that Theorem 4.3.3 is also true for expectation bounded risk measures, which is a generalization of Theorem 4.3.3 to expectation bounded risk measures. Here we just briefly hint why the hedging problem (4.3.4) is not bounded in the presence of Good Deals in this new sense. Suppose that Y^* is a solution to

(4.3.4) *and* $Y \in \mathcal{Y}$ *is a Good Deal. We will show that* $Y^* + Y$ *is a better solution than* Y^**. Indeed, according to sub-additivity of* ρ *and* π

$$\rho(Y^* + Y - X) \leq \rho(Y^* - X) + \rho(Y) < \rho(Y^* - X),$$

$$\pi(Y^* + Y) \leq \pi(Y^*) + \pi(Y) \leq \pi(Y^*) \leq c,$$

$$Y^* + Y \in \mathcal{Y}.$$

This shows that hedging is impossible when there is a Good Deal.

5.3. More About the Existence of Good Deals

In the process of risk assessment (for example capital requirement assessment), two issues are of great importance for the person who assesses the risk: first, she must be well-informed, second, she must be well-equipped. In other words, the ρ*-user on one hand must use an appropriate risk measure and model, and on the other hand must receive and enough (and correct) information. Shortcomings, in each of these two issues may be a source of making wrong decisions, sometimes resulting in risky products, being deemed to be good deals. Here, we use the form "good deal" in the colloquial sense. That's what we can also observe in our theory, when we see how imperfect information may cause risk underestimation and accordingly may produce Good Deals (in the formal sense we have defined). In upcoming discussions, we first focus on the imperfect information, and show how it can produce a Good Deal. That can be an objective for independent research which is outside the scope of this thesis, while we give also some hints to clarify the idea. As for the second issue, being well-equipped, we also give an example at the end of our discussion to show how using Value at Risk as a risk measure may produce a Good Deal.*

5.3.1. Risk Underestimations and Price Underestimations

In this section, we just recall how risk underestimations (in terms of solvency, under-capitalization) and price underestimations would produce Good Deals. According to Theorem 4.3.3, a Good Deal exists when $\mathcal{R} \cap \Delta_\rho \neq \emptyset$*. This means Good Deals will disappear by enlarging the set* Δ_ρ *or* $\mathcal{R}$ *up to touching each other.*

Enlarging Δ_ρ is a financial issue which gives a modified risk measure and is discussed in Section 4.4. However, enlarging $\mathcal{R}$ is an issue in economics. That is because by adding new stochastic discount factors, the equilibrium in the market moves. This is out of the scope of this thesis and it needs to be treated on its own. Since we can rule out Good Deals by enlarging Δ_ρ, which yields a larger risk measure, we interpret this as a modification. This shows how Good Deals are result of the risk underestimation.

Next, we discuss how this underestimation may be may be due to imperfect information.

5.3.2. Imperfect Information

In probability, information can enter into the model via σ-fields. Given two σ-fields $\mathcal{G}$ and $\mathcal{F}$, we always say $\mathcal{F}$ provides more information compared to $\mathcal{G}$ if $\mathcal{G} \subseteq \mathcal{F}$. Here we illustrate with an example how imperfect information may produce Good Deals.

Let m be a stochastic discount factor which gives the pricing rule $\pi(X) = E[mX]$. Consider that m has a continuous distribution function. Let us consider that there is a coherent risk measure ρ such that its associated set Δ_ρ contains m. By Theorem 4.3.3 we know that the couple (ρ, π) does not produce any Good Deal. Now let us consider that a manager who uses the risk measure ρ only has access to the information provided by a finite σ-field $\mathcal{G} = \sigma(\Sigma)$, where Σ is a finite partition of Ω. That means for any position X, the best way the manager can look at X is through $E[X|\mathcal{G}]$. Indeed, the projection of $X \in L^2(\Omega, \mathcal{F})$ to the space $L^2(\Omega, \mathcal{G})$ is $E[X|\mathcal{G}]$. From the manager's point of view, the risk of X is quantified as $\rho(E[X|\mathcal{F}])$. Therefore, we define the following risk measure

$$\rho_\Sigma(X) = \sup_{f \in \Delta_\rho} E[-fE[X|\mathcal{F}]] = \sup_{g \in (\Delta_\rho)_\Sigma} E[-gX], \tag{5.3.1}$$

where $(\Delta_\rho)_\Sigma = \{E[f|\mathcal{F}] | f \in \Delta_\rho\}$. It is quite clear that $\rho_\Sigma(X) = \rho(E[X|\Sigma])$, which shows that ρ_Σ is the real risk measure used by the manager. Let us look at the set $(\Delta_\rho)_\Sigma$. It is clear that this set consists of measures whose distributions are discontinuous while the stochastic discount factor m has continuous distribution. Therefore, it is not contained in $(\Delta_\rho)_\Sigma$. By part 3 of Theorem 4.3.3, it turns out *that a Good Deal must exist. One can argue that the lack of information is due to the manager's information, since the price is something determined in the market and one cannot consider that the stochastic discount factor is $E[m|\mathcal{G}]$. This also means that only the risk measurement is affected by imperfect information.*

5.3.2.1. *Law Invariant Coherent Risk Measures and Good Deals*

In Section 4.3.1 we have seen that how law invariant risk measures can produce Good Deals in a very known models such as Black-Scholes model. In this section we also show that law invariant risk measures are sensitive with respect to imperfect information and complexity. That means if we have less information or less complexity (less added independent random source) in the payoff random variable, the risk assessment might fail to be accurate. In this section we show how using a law invariant risk measure, together with imperfect information or wrong assessment of the payoff, increases the existence of Good Deals.

Imperfect Information. *The most known risk measures such as Value at Risk, Standard Deviation and Expected Shortfall, are law invariant. Also, the classical approach in finance using Expected Utility, is a law invariant approach in assessing risk. Law invariant coherent risk measures on L^∞ are exactly the dilation monotone coherent risk measures (see [27]). As dilation monotone risk measure ρ, we mean a risk measure for which*

$$\rho(E[X|\mathcal{G}]) \leq \rho(X) \tag{5.3.2}$$

for any σ-field $\mathcal{F}$. Indeed, every law invariant convex function on L^∞ is dilation monotone (see [27]). For alternative proofs one could also consult [43] and [50].

Now let us see that for a law invariant coherent risk measure ρ, we have

$$\rho(E[X|\mathcal{G}]) + \pi(X) \leq \rho(X) + \pi(X).$$

According to Theorems 4.3.3 and 5.3.2 one can easily see that it is more likely that in imperfect information, using a law invariant risk measure, one would come up with a Good Deal.

Complexity and Good Deals. *By complexity, we mean a concept which symbolizes the amount of the fluctuations of a random variable. In mathematical terms, we say that a random variable Y is more complex than X, if there exists another random variable ϵ, independent of X, with zero mean, such that $Y \stackrel{d}{=} X+\epsilon$. The random variable ϵ can be looked as a source of information which is invisible for the person who detects X. The law invariant risk measures are sensitive about complexity because, $\rho(X) \leq \rho(Y)$. Actually, there exists $X' \sim X$ such that $E[Y|X'] = X'$ from which one has $\rho(X) = \rho(X') \leq \rho(Y)$. For a moment we take $\mathcal{Y}$ as subspace of L^p. If the real payoff is a function of $\mathcal{Y}+\epsilon$ instead $\mathcal{Y}+\epsilon$, while the ρ-user only observe $\mathcal{Y}$, then the risk of producing a Good Deal would grow.*

5.3.3. Robust Risk Measures and Good Deals

In this last section, we briefly discuss the relation between robustness and Good Deals. This is to show that how tools we have chosen to use may be a source of producing Good Deals. This is also important because of the critique made in [29] about the un-robustness of coherent risk measures. They introduced a robust family of risk measures instead. We found that this family of risk measures in a perfect market always produce Good Deals, which shows a dilemma in risk assessment, while a stable assessment is possible only if we accept the existence of Good Deals.

In [29], it has been shown that a law-invariant coherent risk measure is not robust, as has been also discussed in Chapter 3 and in Section 3.5.1. Indeed, the authors in [29], after showing that law invariant risk measures are not robust, study the robustness of the following family of risk measures:

$$\rho(X) = \int_0^1 VaR_\alpha(X)\phi(\alpha)d\alpha, \tag{5.3.3}$$

for a density ϕ. They proved that the risk measure ρ in (5.3.3) *is robust if and only if there exists $\beta > 0$ such that* $\operatorname{supp}(\phi) \subseteq [\beta, 1-\beta]$ *(see also Section 3.5.1).*

In [8], we found that the family (5.3.3) *of risk measures in a perfect market always produce Good Deals. We quote the following theorem from [8]:*

Theorem 5.3.1. *Every robust risk ρ in a perfect market with pricing rule π generates Good Deals.*

The reason why this happens is easy to see. This is because ϕ, for robust risk measures, is zero in a neighborhood around 0 *and* 1*, which shows ρ ignores information that appear in the tails. This makes one able to construct a Good Deal based on the unseen information by ρ in the tails.*

To justify how by using a risk measure like 5.3.3, a Good Deal is produced, we give the following simple example:

Example*: Suppose that an event is going to happen tomorrow, with probability* 1 *percent. We construct a simple security, which penalizes the security-owner for 100 dollar if the* 1*-percent event happens, and nothing otherwise. The price of this security is simply $-100 \times \frac{1}{100} + 0 \times \frac{99}{100} = -1$ dollar. We denote this security with X i.e., $\pi(X) = -1$. It says that if someone trades X, she should be payed* 1 *dollar, or if she is risk-averse by more than one dollar. Therefore, let us consider that the price is $-c$ dollar, less than -1, but also not very far (at least we consider $-1 > -c > -20$). Consider there is a financial practitioner who is endowed with the risk measure $\mathrm{VaR}_{0.05}$. From her point of view, the risk associated to this security is zero (because she is not able to see events in the* 5*-percent of the tail) while she is paid $-c > 0$. This shows how $\mathrm{VaR}_{0.05}$ ignores information in the* 5 *percent in the tail and also shows how the security constructed based on the* 1*-percent event would produce a Good Deal ($\mathrm{VaR}_{0.05}(X) + \pi(X) = 0 - c < 0$).*

Now let us see what happens if one takes $\mathrm{CVaR}_{0.05}$ in lieu of $\mathrm{CVaR}_{0.05}$. It is clear that $\mathrm{CVaR}_{0.05}(X) = \frac{1}{0.05}\int_0^{0.01} 100ds = 20$ dollar, which is assessed quite riskier than before. On the other hand also $\mathrm{CVaR}_{0.05}(X)+\pi(X) = 20-c > 0$ and then by Theorem 4.3.3 this means that using $\mathrm{CVaR}_{0.05}$, X is not a Good Deal.

CONCLUSION

The final utility of the analyzes presented in this thesis can be found in capital requirement applications. In particular, one relevant element that arises within our analysis is the fact that underestimation in capital requirement might destabilize the market. These applications have been extensively discussed within the thesis, here we briefly mention again the three specific achievements in this thesis.

In Chapter 2, we have shown that the Lebesgue property of a risk measure on bounded càdlàg processes is equivalent to the Lebesgue property of its static version. We also used the results of Chapter 2 to solve the problem of capital allocation in this setting.

In Chapter 3, we have shown that how the concept of natural risk statistics could be extended to the space of infinite sequences. We also have shown that how this extension could be used to derive a consistence family of natural risk statistics for any dimension.

In Chapter 4, we have studied the situation when a market is destabilized in the presence of Good Deals. We have shown how this situation can be recovered with modifying a risk measure to a larger one. We also proposed two different ways of modifying risk, based on minimal risk spread criteria and maximal Global/Local preference ratio.

In this work, we developed different ideas but there are much more avenues that remained unexplored. We believe that the work we started in this thesis led the way to new interesting directions that are yet to be investigated. These new directions are briefly discussed in Chapter 5 where we tried to connect all ideas in previous chapters and by generalizing the definition of Good Deals. In particular, *we discussed the relevant problem of how Good Deals arise and how they might be the product of lack of information.*

Bibliography

[1] *S. Ahmed, D. Filipović, and G. Svindland. A note on natural risk statistics.* Oper. Res. Lett.*, 36(6):662–664, 2008.*

[2] *C. D. Aliprantis and K. C. Border.* Infinite dimensional analysis. *Springer, Berlin, third edition, 2006. A hitchhiker's guide.*

[3] *C. D. Aliprantis and O. Burkinshaw.* Locally solid Riesz spaces. *Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1978. Pure and Applied Mathematics, Vol. 76.*

[4] *K. J. Arrow.* Social Choice and Individual Values. *Cowles Commission Monograph No. 12. John Wiley & Sons Inc., New York, N. Y., 1951.*

[5] *P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk.* Math. Finance*, 1999.*

[6] *S. Asmussen.* Ruin probabilities*, volume 2 of* Advanced Series on Statistical Science & Applied Probability. *World Scientific Publishing Co. Inc., River Edge, NJ, 2000.*

[7] *H. Assa. Lebesgue property of risk measures for bounded càdlàg processes and applications.* `http://www.gerad.ca/fichiers/cahiers/G-2009-16.pdf`.

[8] *H. Assa. A robustness dilemma in risk measuring.* preprint.

[9] *H. Assa and A. Balbás. Good deals: Compatible extension of risk and pricing rule.* `http://www.gerad.ca/fichiers/cahiers/G-2009-54.pdf`.

[10] *H. Assa and M. Morales. Risk measures on the space of infinite sequences.* Math. Financ. Econ.*, 2(4):253–275, 2010.*

[11] *J.-P. Aubin. Coeur et équilibres des jeux flous sans paiements latéraux.* C. R. Acad. Sci. Paris Sér. A*, 279, 1974.*

[12] *A. Balbás, B. Balbás, and A. Heras. Optimal reinsurance with general risk measures.* Insurance Math. Econom.*, 44(3):374–384, 2009.*

[13] *A. Balbás and R. Balbás. Compatibility between pricing rules and risk measures: The ccvar. pages 251–264, 2009.*

[14] *A. Balbás, R. Balbás, and J. Garrido. Extending pricing rules with general risk functions.* European J. Oper. Res.*, 201(1):23–33, 2010.*

[15] *A. Balbás, R. Balbás, and S. Mayoral. Portfolio choice and optimal hedging with general risk functions: a simplex-like algorithm.* European J. Oper. Res.*, 192(2):603–620, 2009.*

[16] *A. Balbás and P. Jiménez Guerra. Generalized vector risk functions.* `http://e-archivo.uc3m.es:8080/bitstream/10016/514/1/wb066721.pdf`.

[17] *L. J. Billera and D. C. Heath. Allocation of shared costs: a set of axioms yielding a unique procedure.* Math. Oper. Res.*, 7(1):32–39, 1982.*

[18] *A. Buch and G. Dorfleitner. Coherent risk measures, coherent capital allocations and the gradient allocation principle.* Insurance Math. Econom.*, 42(1):235–242, 2008.*

[19] *A. Cerný and S. Hodges. Coherent risk measures on general probability spaces. In* Mathematical Finance, Bachelier Congress 2000*, pages 175–202. Springer Verlag, Berlin, 2002.*

[20] *T. Chan. Pricing contingent claims on stocks driven by Lévy processes.* Ann. Appl. Probab.*, 9(2):504–528, 1999.*

[21] *P. Cheridito, F. Delbaen, and M. Kupper. Coherent and convex monetary risk measures for bounded càdlàg processes.* Stochastic Process. Appl.*, 112(1):1–22, 2004.*

[22] *P. Cheridito, F. Delbaen, and M. Kupper. Coherent and convex monetary risk measures for unbounded càdlàg processes.* Finance Stoch.*, 9(3):369–387, 2005.*

[23] *P. Cheridito, F. Delbaen, and M. Kupper. Dynamic monetary risk measures for bounded discrete-time processes.* Electron. J. Probab.*, 11:no. 3, 57–106 (electronic), 2006.*

[24] *A. Cherny. Capital allocation and risk contribution with discrete-time coherent risk.* Math. Fin., *19(1):13–40, 2009.*

[25] *A. S. Cherny. Equilibrium with coherent risk. Quantitative Finance Papers math/0605051, arXiv.org, May 2006.*

[26] *A. S. Cherny. Pricing and hedging European options with discrete-time coherent risk.* Finance Stoch., *11(4):537–569, 2007.*

[27] *A. S. Cherny and P. G. Grigoriev. Dilatation monotone risk measures are law invariant.* Finance Stoch., *11(2):291–298, 2007.*

[28] *J. H. Cochrane and J. Saa-Requejo. Beyond arbitrage: Good deal asset price bounds in incomplete markets.*

[29] *R. Cont, R. Deguest, and G. Scandolo. Robustness and sensitivity analysis of risk measurement procedures.* Quantitative Finance, *10(6):593–606, 2010.*

[30] *G. Debreu.* Theory of value: an axiomatic analysis of economic equilibrium. *Cowles Foundation for Research in Economics at Yale University, Monograph 17. John Wiley & Sons Inc., New York, 1959.*

[31] *F. Delbaen. Risk measures or measures that describe risk?* `http://www.math.ethz.ch/~delbaen/ftp/preprints/Risk-executives.pdf`.

[32] *F. Delbaen.* Coherent risk measures. *Cattedra Galileiana. [Galileo Chair]. Scuola Normale Superiore, Classe di Scienze, Pisa, 2000.*

[33] *F. Delbaen. Coherent risk measures on general probability spaces. In* Advances in finance and stochastics, *pages 1–37. Springer, Berlin, 2002.*

[34] *C. Dellacherie and P.-A. Meyer.* Probabilités et potentiel. Chapitres V à VIII, *volume 1385 of* Actualités Scientifiques et Industrielles [Current Scientific and Industrial Topics]. *Hermann, Paris, revised edition, 1980. Théorie des martingales. [Martingale theory].*

[35] *M. Denault. Coherent allocation of risk capital.* Journal of Risk, *4:1–34, 2001.*

[36] *I. Ekeland and R. Témam.* Convex analysis and variational problems, *volume 28 of* Classics in Applied Mathematics. *Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, english edition, 1999. Translated from the French.*

[37] *L. Epstein and T. Wang. Intertemporal asset pricing under knightian uncertainty.* Econometrica, Econometric Society*, 62(2):283–322, 1994.*

[38] *D. Filipović, M. Kupper, and N. Vogelpoth. Separation and duality in locally L^0-convex modules.* J. Funct. Anal.*, 256(12):3996–4029, 2009.*

[39] *D. Filipovic and G. Svindland. The canonical model space for law-invariant convex risk measures is L^1.* `http://sfi.epfl.ch/page85019.html`*, 2008.*

[40] *T. Fischer. Risk capital allocation by coherent risk measures based on one-sided moments.* Insurance Math. Econom.*, 32(1):135–146, 2003.*

[41] *H. Föllmer and A. Schied. Convex measures of risk and trading constraints.* Finance Stoch.*, 6(4):429–447, 2002.*

[42] *H. Föllmer and A. Schied. Robust preferences and convex measures of risk. In* Advances in finance and stochastics*, pages 39–56. Springer, Berlin, 2002.*

[43] *H. Föllmer and A. Schied.* Stochastic finance*, volume 27 of* de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, extended edition, 2004. An introduction in discrete time.*

[44] *A. Grothendieck.* Topological vector spaces*. Gordon and Breach Science Publishers, New York, 1973. Translated from the French by Orlando Chaljub, Notes on Mathematics and its Applications.*

[45] *M. Gundlach and F. Lehrbass, editors.* CreditRisk$^+$ in the banking industry*. Springer Finance. Springer-Verlag, Berlin, 2004.*

[46] *A. H. Hamel, F. Heyde, and M. Höhne. Set-valued measures of risk.* Reports on Optimization and Stochastics, Martin-Luther-University Halle-Wittenberg*, (15), 2007.*

[47] *F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel.* Robust statistics*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. The approach based on influence functions.*

[48] *G. Jehle and P. Reny.* Advanced microeconomic theory*. Addison-Wesley series in economics. Addison-Wesley, 2001.*

[49] *E. Jouini, M. Meddeb, and N. Touzi. Vector-valued coherent risk measures.* Finance Stoch.*, 8(4):531–552, 2004.*

[50] *E. Jouini, W. Schachermayer, and N. Touzi. Law invariant risk measures have the Fatou property. In* Advances in mathematical economics. Vol. 9*, volume 9 of* Adv. Math. Econ.*, pages 49–71. Springer, Tokyo, 2006.*

[51] *D. Kannan and V. Lakshmikantham.* Handbook of stochastic analysis and applications*, volume 163 of* Statistics: Textbooks and Monographs*. Marcel Dekker Inc., New York, 2002.*

[52] *C. Klüppelberg, A. E. Kyprianou, and R. A. Maller. Ruin probabilities and overshoots for general Lévy insurance risk processes.* Ann. Appl. Probab.*, 14(4):1766–1801, 2004.*

[53] *S. Kou, C. Heyde, and X. Peng. What is a good risk measure: Bridging the gaps between data, coherent risk measures, and insurance risk.* `http://www.cfe.columbia.edu/pdf-files/Heyde_Kou_07_07.pdf`*.*

[54] *R. Mansini, W. Ogryczak, and M. G. Speranza. Conditional value at risk and related linear programming models for portfolio optimization.* Ann. Oper. Res.*, 152:227–256, 2007.*

[55] *A. J. McNeil, R. Frey, and P. Embrechts.* Quantitative risk management*. Princeton Series in Finance. Princeton University Press, Princeton, NJ, 2005. Concepts, techniques and tools.*

[56] *M. Morales and W. Schoutens. A risk model driven by Lévy processes.* Appl. Stoch. Models Bus. Ind.*, 19(2):147–167, 2003.*

[57] *Y. Nakano. Efficient hedging with coherent risk measure.* J. Math. Anal. Appl.*, 293(1):345–354, 2004.*

[58] *R. T. Rockafellar, S. Uryasev, and M. Zabarankin. Generalized deviations in risk analysis.* Finance Stoch.*, 10(1):51–74, 2006.*

[59] *A. Ruszczyński and A. Shapiro. Optimization of convex risk functions.* Math. Oper. Res.*, 31(3):433–452, 2006.*

[60] *S.Alexander, T. Coleman, and Y.Li. Minimizing cvar and var for a portfolio of derivatives.* Journal of Banking & Finance*, 30(2):583 – 605, 2006. Risk Management and Optimization in Finance.*

[61] *A. Schied. Optimal investments for risk- and ambiguity-averse preferences: a duality approach.* Finance Stoch.*, 11(1):107–129, 2007.*

[62] *J. Staum. Fundamental theorems of asset pricing for good deal bounds.* Math. Finance*, 14(2):141–161, 2004.*

[63] *D. Tasche. Expected shortfall and beyond.* Journal of Banking and Finance*, 26(7):1519–1533, 2002.*

[64] *D. Tasche. Capital allocation for credit portfolios with kernel estimators.* Quant. Finance*, 9(5):581–595, 2009.*

[65] *S. Wang. Insurance pricing and increased limits ratemaking by proportional hazards transforms.* Insurance Math. Econom.*, 17(1):43–54, 1995.*

[66] *S. S. Wang, V. R. Young, and H. H. Panjer. Axiomatic characterization of insurance prices.* Insurance Math. Econom.*, 21(2):173–183, 1997.*

[67] *C. Zalinescu.* Convex analysis in general vector spaces*. World Scientific Press Co., 2002.*