A Study on Dilatation Monotone Maps and the Haezendonck-Goovaerts Risk Measure

by

Massoomeh Rahsepar
Master of Science, Ryerson University, 2011
Master of Science, Isfahan University of Technology, 2004
Bachelor of Science, Mohaghegh Ardebily University, 2000

A dissertation
presented to Ryerson University
in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy (Ph.D.)
in the Program of
Mathematical Modelling and Methods

Toronto, Ontario, Canada, 2021

Author's Declaration For Electronic Submission Of A Dissertation

I hereby declare that I am the sole author of this dissertation. This is a true copy of the dissertation, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this dissertation to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this dissertation by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my dissertation may be made electronically available to the public.

A Study on Dilatation Monotone Maps and the Haezendonck-Goovaerts Risk Measure

Massoomeh Rahsepar
Doctor of Philosophy (Ph.D.)
Mathematical Modelling and Methods
Ryerson University, 2021

Abstract

Risk measures have been studied for several decades in the financial literature. Mathematically, a risk measure is a mapping from a class of random variables defined on some space $\mathcal{X}$ to the (extended) real line. It is difficult to give a reliable assessment of financial risk without a suitable model space $\mathcal{X}$ where the financial position lies there.

In this work, we assume the domain $\mathcal{X}$ for the risk measure $\rho : \mathcal{X} \to (-\infty, \infty]$ is a subset of L^1 that contains the space of simple random variables $\mathcal{L}$ and has the Fatou property. We focus on dilatation monotonicity as desirable preferences of the decision-maker over risks. We discuss continuity properties of dilatation monotone risk measures on $\mathcal{X}$. We prove that on a non-atomic probability space, every dilatation monotone convex risk measure on $\mathcal{X}$ can extend uniquely to a $\sigma(L^1, \mathcal{L})$-lower semicontinuous dilatation monotone convex risk measure. Our findings complement extension and continuity results for (quasi)convex law-invariant functionals.

In the second part of this thesis, we overview various properties of the Haezendonck–Goovaerts risk measure as a dilatation monotone risk measure. In the special case $\Phi(x) = x^2$, we develop an explicit expression for the Haezendonck–Goovaerts risk measure when the risk variable X follows some specific distributions and propose an empirical algorithm to estimate this risk measure. Moreover, we investigate the performance of some members of this class of risk measures on real data as a possible application of the Haezendonck–Goovaerts risk measure.

Key Words: Dilatation Monotone, Extension of Risk Measures, Fatou Property, Haezendonck–Goovaerts Risk Measure, Average Value at Risk, Optimization, Empirical Algorithm.

Acknowledgments

There are many people to whom I owe thanks over the past several years as I worked toward my Ph.D. degree. Without their help and support, it would not be possible for me to present this thesis. First and foremost, I would like to acknowledge special thanks to my supervisor, Dr. Xanthos, who has been a great mentor and friend to me for the whole period of my Ph.D. study. I feel blessed and precious for having him throughout this journey.

I would like to thank all my commitee/examination members, Dr. Emanuela Rosazza Gianin, Dr. Yuanshun Li, Dr. You Liang and Dr. Niushan Gao, for agreeing to serve on my Ph.D. committee, and their time and consideration when reading this thesis.

Finally, I would like to warmly thank my family, especially my husband, for all his support, patience, and company, and my parents and my sister who have always supported me through any endeavours I chose to take on, no matter how big or small, and I am very grateful to have such an amazing family.

”Do not judge me by my successes; judge me by how many times I fell down and got back up again.”

— Nelson Mandela

Dedication

“ Call it a clan, call it a network, call it a tribe, call it a family: Whatever you call it, whoever you are, you need one. ” -–Jane Howard

To my husband, Ali, and my daughters, Elina and Romina.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation from Financial Mathematics

Monetary Risk Measure

Our modern society relies on well-functioning banking and insurance systems. In the financial world, facing long-term future uncertainties places financial institutions into various risks, which are not easy to quantify and are even more challenging to control. Any reckless speculative activities can cause crises or collapse in the financial system. The global financial crisis (2007-2008) is one example of excessive risk-taking by banks.

The need to quantify and manage risks is of fundamental importance in finance. Great effort has been invested in developing risk measurement techniques that can bridge the gap between financial institutions and practitioners, with the aim of achieving reliable and easy to implement methodologies for measuring risks. In financial mathematics, a risk measure is used to determine the amount of capital (assets, currency, etc.) to be held in reserve to keep financial institutions such as banks and insurance companies solvent in a crisis and cover unexpected losses. This amount of necessary capital depends on the risk attached to a particular bank's assets or financial positions.

After a couple of decades of introducing the theoretical framework of (coherent) risk measures by Artzner et al. (1999), many questions still originate. One of the issues that an individual can face is choosing a suitable model space $\mathcal{X}$ for the financial positions. Risk measures are often considered as maps on bounded random variables. However, this desirable space is clearly too small to capture the actual risks. A natural way in the direction beyond bounded random variables is to pick up a particular space and then reconstruct a whole theory with careful analysis of the new space's structure, like the L^p, Orlicz, Orlicz hearts or even L^0 spaces. On the other hand, it seems more efficient to extend a risk function ρ defines on the set $\mathcal{X}$ to some larger space while still preserving its desirable properties such as convexity, subadditivity, cash-invariant, monotonicity (regular or even order). In this sense, Filipović and Svindland (2012) have shown that any law invariant convex risk measure with the Fatou property on the space of bounded random variable is uniquely extended to the L^1 space preserving the Fatou property. In this dissertation, one of our main focuses is to extend risk maps ρ that initially define on some suitable space $\mathcal{X}$ to the L^1 space.

Putting Risk Aversion into Risk Measures

It has been common to argue that individuals tend to display an aversion to the taking of risks. Risk aversion is a fundamental parameter to determine individuals' investment decisions, and it, in turn, is an explanation for many observed phenomena in the financial world. Therefore, when one studies risk measures, she/he should also model the behaviour of decision-makers over risks.

There is a long tradition in the risk management field to use some sort of preference ordering such as stochastic ordering or convex ordering to formulate some favourable properties of risk maps. The risk aversion that we focus on in this thesis is called dilatation monotonicity. Dilatation monotonicity is reflected in the general belief that balancing out should never increase the involved risk. Namely for two positions X and $Y = \mathbb{E}[X|.]$, If a risk measure ρ has the dilatation monotonicity property, then the capital requirement for position Y should be smaller than the capital requirement for position X (i.e. $\rho(Y) \leq \rho(X)$). Consequently, the financial institution has an incentive to choose Y over X.

A Demand for Tail Sensitive Risk Measures

The (market) risk model was built upon assumptions that were not reflective of the real world, especially in stressed financial markets. Financial turbulence such as the global financial crisis or the recent Covid-19 recession dramatically demonstrates that risk management tools had often been chosen wrongly. Recall that tail risk events underestimate possible losses in stressed conditions on the quantile risk measure that often used in practice. As such, it may be plausible to consider risk measures that are more sensitive to tail events.

An interesting example of such a risk measure is the Haezendonck–Goovaerts risk measure, which is a coherent risk measure with the dilatation monotonicity property. The Haezendonck–Goovaerts risk measure is introduced by J. Haezendonck and M. Goovaerts (1982), and it is defined via a convex Young function and a confidence level parameter $\alpha \in (0, 1)$. Roughly speaking, this measure quantifies risk using the higher-order moments by putting more weight on the right tail. In this thesis, we consider the Haezendonck–Goovaerts risk measure with a power Young function $\Phi(x) = x^p$, $p \geq 1$. The Haezendonck–Goovaerts risk measure with a power Young function is differentiable. Thus, this coherent dilatation monotone risk measure becomes a favourable option for our study. The best-known example of the risk measure in this class is the Average Value at Risk corresponds to the most straightforward Young function $\Phi(x) = x$.

1.2 Thesis Outline

This thesis can be divided into two parts. Each part is written to be read independently with self-contained notations and definitions. If reference is made to another chapter, the connection is clearly explained. The outline of the thesis is as follows.

In Chapter 2 which is based on section 2, [59], we provide a sound approach towards the study of dilatation monotone risk maps $\rho : \mathcal{X} \to (-\infty, \infty]$ where $\mathcal{X}$ is a subset of L^1 and contains the space of simple random variables $\mathcal{L}$. We show that how one can construct a finite partition π, such that any $X \in L^1$ can be approximated by a conditional expectation with respect to finitely generated σ-algebra $\sigma(\pi)$, namely $\mathbb{E}[X|\sigma(\pi)]$. This result is a core for many results in this Chapter which is independent interest. Moreover, base on our framework, we prove that on a non-atomic probability space, dilatation monotonicity property coupled with the Fatou property implies that the underlying risk measure ρ is lower semicontinuous with respect to the relative $\sigma(L^1, \mathcal{L})$-topology. We also show that ρ extends uniquely to a $\sigma(L^1, \mathcal{L})$-lower semicontinuous and dilatation monotone functional $\overline{\rho} : L^1 \to (-\infty, \infty]$. While most results on the extension theory of law-invariant risk measures require convexity of risk measures, the novelty of our result is that we do not impose any quasiconvexity assumptions on the underlying risk measures. Finally, we conclude this Chapter by investigating the norm continuity of cash-invariant hulls. An interesting example of this cash-invariant hull class is the Haezendonck–Goovaerts risk measure. This risk measure is the subject of our study for the Next Chapters

In the second part of this dissertation, we study the well-known Haezendonck-Goovaerts risk measure. The focus of Chapter 3 is the Haezendonck-Goovaerts risk measure which is defined with a power Young function $\Phi(x) = x^p,\ p \geq 1$. In this Chapter, we analyze the properties of Orlicz quantiles. One of main contribution is theorem 40 where we prove that with $\Phi(x) = x^2$, for any given $\alpha \in (0,1)$, the Orlicz quantile for the Haezendonck-Goovaerts risk measure is bounded by the α-quantile which is the Orlicz quantile of the Average Value at Risk. By considering the cases where risk variable X follows normal distributions, we investigate an analytical or numerical expression for the Haezendonck-Goovaerts risk measure. Besides, we develop an empirical method to implement the Haezendonck-Goovaerts risk measure when $\Phi(x) = x^2$ (Theorem 54) .

The central point of Chapter 4 is to study a potential application of this class of risk measures in actual data. First, we provide an analytical study by applying our suggested approximation method to calculate this risk measure when $p = 2$ using different distributions such as uniform, exponential and normal distributions. To this end, we apply the empirical method by allowing the data to speak for themselves in the sense that the future values model should be based on information available in the given historical data sample. In this case study, we investigate the theory that higher values of $p > 1$ may provide a better indicator of measuring risk than the Average Value at Risk. We suggest that when we employ this risk measure to a sample with unknown underlying distributions, the empirical method could provide an efficient way to estimate the Haezendonck-Goovaerts risk measure. This data-driven study indicates the advantages of the proposed measure compared to the Average Value at Risk, especially in times of significant turbulence in riskier scenarios in which a proper risk measure is vital.

Chapter 2

A Study on Dilatation Monotone Maps

Over the past few years, risk management and risk measurement have been active fields of study with numerous unsolved problems and issues to address. One of these issues is selecting a suitable model space $\mathcal{X}$ for risk measures. In Section 2.2, we review some of the model spaces that investigated in the literature. Section 2.3.2 contains crucial results, which help us to study dilatation maps. Then, we study the continuity properties of dilatation monotone maps $\rho : \mathcal{X} \to [-\infty, \infty]$. We elaborate on the extensions theorem's existence and uniqueness in Section 2.4 where we extend dilatation risk measures with the Fatou property from not a necessarily vector space $\mathcal{X}$ to the L^1 space. We conclude this Chapter by investigating the continuity of monotone and cash-invariant hulls.

2.1 Risk Functions and Risk Measures

2.1.1 Introduction

Risk is known as a particular case of uncertainty. In our everyday lives, we always make important and risky decisions. Most of these decisions are made without knowing their consequences with certainty because they depend on events that may or may not occur. Such decision problems arise, for instance, in financial and insurance markets as well. Financial institutions face a variety of risks, which are not easy to quantify. While the basic understanding of a risk measure has been well-accepted by academics and practitioners for many decades, the mathematical formulation of risk measure is a new topic in the mathematical finance literature. An essential contribution in this direction is given in the seminal 1999 paper by Artzner et al. [5]. In the following few paragraphs, we will introduce the reader to some classical concepts on risk measures. For more details on this topic, we refer the reader to [31], Chapter 4.

2.1.2 Risk Measures Base on Capital Requirements

In general, a risk measure should quantify the downside risk of a financial position in monetary units. We assume that the uncertainty that a decision-maker faces can be represented by a non-empty set of states denoted as Ω. This set may be finite or infinite. Associated with Ω is the set of events, taken to be a σ-algebra of subsets of Ω, denoted by $\mathcal{F}$. The uncertain outcomes of a financial position (such as a portfolio of stocks) can be described by a function $X : \Omega \to \mathbb{R}$, where $X(\omega)$ is the value of the position at the end of the trading period if the state $\omega \in \Omega$ occurs. A risk measure ρ is defined on a class $\mathcal{X}$ of financial positions quantifying the risk (uncertain outcome) of X by some number, possibly extended to the real line. Note that any risk measure has a natural domain $\mathcal{X} \subset L^0$, where $L^0 := L^0(\Omega, \mathcal{F}, \mathbb{P})$ is the class of all measurable elements on $(\Omega, \mathcal{F}, \mathbb{P})$.

Let us list some properties that have been proposed as natural requirements for adequate risk measures. Let $\mathcal{X}$ be a suitable model space, then the (risk) map ρ may fulfill the following properties:

(A1) Monotonicity: For any $X, Y \in \mathcal{X}$ with $X \geq Y,\ \rho(X) \leq \rho(Y)$.

Monotonicity property is consistent with the fact that higher capitals bear less risks in the monetary agency's view.

(A2) Cash-invariance: For any $X \in \mathcal{X}$ and $c \in \mathbb{R}$ with $X + c \in \mathcal{X},\ \rho(X + c) = \rho(X) - c$.

The cash-invariance means that adding a constant amount of c to X should reduce the capital requirement for X by the amount of c. Note that, for an unacceptable portfolio X, adding the amount $\rho(X)$ makes the position acceptable.

A risk measure satisfying the properties of cash-invariance and monotonicity is called a (monetary) risk measure which is the minimum capital should be added to the position X to make it acceptable from the point of view of a supervising agency. We call a monetary risk measure convex if it also satisfies

(A3) Convexity: For any $X, Y \in \mathcal{X}$, and $\lambda \in (0,1)$ such that $\lambda X + (1-\lambda)Y \in \mathcal{X}$ then

$$\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y).$$

A convex monetary risk measure is called coherent if it satisfies

(A4) Positive homogeneity: For any $X \in \mathcal{X}$ and $\lambda \geq 0$, $\rho(\lambda X) = \lambda\rho(X)$.

Under (A3), the capital requirement for the convex combination of two positions does not exceed the convex combination of the individual capital requirement. So the axiom of convexity gives a precise meaning that diversification should not increase the risk. Moreover, if (A4) holds, then the capital requirement scale linearly when the net worth is multiplied with a non-negative constant.

2.1.3 Risk Measures Based on Premium Principles

Let us assume the point of the risk management of an insurance company. In the literature, it is well-known that there is a close connection between risk measures in mathematical finance and the actuarial on premium principles. The premium calculation principle is one of the main objectives of study for actuaries. In insurance terminology, a premium is a payment that policyholders make to obtain protection from their risks. A premium principle is a rule for assigning premiums to insurance risks. In actuarial applications, they often work with loss distributions for insurance products. Additionally, it is usually relevant to consider in insurance contexts that the loss X is non-negative.

Definition 1. *A risk measure $\pi : \mathcal{X} \to \mathbb{R}$ is called an insurance (a reinsurance) premium principle and defines as*

$$\pi(X) := \rho(-X).$$

Essentially, the monotonicity condition and the cash-invariance property change for the premium principle π while other properties remain the same.

(B1) Monotonicity: For any $X, Y \in \mathcal{X}$ with $X \leq Y$, $\pi(X) \leq \pi(Y)$.

(B2) Cash-invariance: For any $X \in \mathcal{X}$ and $c \in \mathbb{R}$ with $X + c \in \mathcal{X}$, $\pi(X + c) = \pi(X) + c$.

The monotonicity says a portfolio with more potential loss is riskier than the other. If (B2) holds, adding or withdrawing a constant amount c to or from a loss X results in decreasing or increasing the risk by precisely the amount c.

Two of the most popular risk measures present in the literature are the Value at Risk ($V@R$) and the Average Value at Risk [1].

[1]Different authors have proposed similar concepts using different names, the Expected Shortfall [2] , the Conditional Value at Risk [62], the Tail Conditional Expectation [5], the Average Value at Risk [31].

Definition 2. *Let X be a loss of a portfolio. For a fixed $\alpha \in (0,1)$, the Value-at-Risk is the α-quantile of X*

$$\mathrm{V@R}_\alpha(X) := F^{-1}(\alpha).$$

Definition 3. *For a portfolio loss X, satisfying $\mathbb{E}[|X|] < \infty$, the Average Value-at-Risk at the confidence level $\alpha \in (0,1)$ is given by*

$$\mathrm{AV@R}_\alpha(X) := \frac{1}{1-\alpha}\int_\alpha^1 F_X^{-1}(u)du.$$

In accord with the premium principles definition, in the actuarial literature, some axiomatic approaches to risk measures based on insurance premium principles have been proposed; one is the Orlicz premium principle (Haezendonck and Goovaerts (1982)). In the Next Chapter, we will study further on the Orlicz premium principle or the Haezendonck-Goovaerts risk measure.

2.2 Model Assumptions

Measuring risk for a financial position is a complicated process. The risk could arise because of potential errors in the models or inappropriate models assumption. The errors and inaccuracies can cause considerable monetary losses, poor organizational decision-making, and institutional reputation damage. Particularly, in the aftermath of the global financial crisis, it is realized that the mathematical models for financial markets carry a substantial amount of the model risk. Recall that one of the main factors to have a reliable financial risk assessment is depends on a suitable model space $\mathcal{X}$ for financial positions. The two most prominent approaches for the model space assumption that can be found in the literature are as the following.

2.2.1 Model Assumption: A General Approach

After two decades of introducing the first axiomatic way of defining risk measures; still, there is an open-end discussion about the right model assumption. One of the primary concerns in this regard is selecting an ideal domain (space of random variables) of the risk maps ρ.

Many scholars argue that the model space should not depend on any particular risk measure and it should include as many risk measures as possible, while some believe that the model space should determine by choice of some specific risk measures. We have to point out that the choice of the domain space $\mathcal{X}$ leads to different representation results. For instance, in most practical problems, such as optimization, features such as finiteness and continuity of functions are crucially important. These features are linked directly to the domain on which the risk functions are defined. Hence, establishing a natural space for the domain of risk maps is an important task.

The most popular domain space for risk measures in the literature is the space of bounded random variables $\mathcal{X} = L^\infty$; see for example [23, 24, 30, 49, 51] and the references therein. Despite well-established theoretical investigations in this model space. However, the L^∞ space is obviously too small to capture the actual risks. Note that financial risks are typically unbounded (e.g. normal distributed random variables are not contained in the L^∞ space) and even heavy-tailed in actuarial applications. This fact becomes a key direction toward the study of risk measures beyond bounded random variables. Therefore, considering $\mathcal{X} = L^p$, $p \in [1, \infty)$ as the domain of risk measures seems quite natural ([28, 43, 44]). Nowadays, the increasing use of heavily-tailed distributions in the risk modelling has led to the generalization of the domain space from the L^p spaces to the space where $\mathcal{X}$ could be Orlicz spaces, Orlicz hearts and other rearrangement invariant spaces (check [11, 12, 20, 52] for more detail).

2.2.2 Model Assumption: An Extension Approach

As already mentioned, when we are dealing with possibly unbounded financial positions X, the L^∞ space is not suitable. Another direction towards dealing with unbounded risks, which one can explore, is to extend risk measures originally defined on the L^∞ space to some larger space.

The concern in this approach is how far a risk measure initially defined on some space $\mathcal{X}$ can be extended in such a way to preserve some of its properties or, roughly speaking, how far we can go to keep a good risk measure to be good again. For example, It is well-known that any cash-invariance risk measure on L^∞ is automatically finite-valued and (Lipschitz) continuous. In parallel, cash invariant risk measures on L^p,

$p \in [1, \infty)$, do not need not to be either finite-valued or continuous. Therefore a natural question is how one can extend a risk measure ρ such that the extended risk measure $\bar{\rho}$ preserves most of the properties of ρ, most importantly, continuity!.

The problem of extending the domain of risk measures to a larger domain has received significant attention in the mathematical finance literature as a solution to deal with a suitable model assumption (see, e.g. [17,29,32,45,52,67]). In this regard, the celebrated work [29], asserts that if $\mathcal{X}$ is some L^p-space for $1 \leq p \leq \infty$, then any law-invariant convex risk measure $\rho : L^p \to (-\infty, \infty]$ with the Fatou property can be extended uniquely to a law-invariant and $\|\cdot\|_1$-lower semicontinuous convex risk measure $\overline{\rho} : L^1 \to (-\infty, \infty]$.

To generalized these extension results, in [32], the authors show that a similar extension result still holds in the setting of a general Orlicz space if one replaces norm-lower semicontinuity by the Fatou property. In [17, 52], the authors utilized the extension property to show that a wide class of optimal risk sharing problems can be solved on any rearrangement invariant commodity space if and only if the problem is solvable on the level of all integrable random variables. In our study, under our model assumption $\mathcal{X}$, we show that dilatation monotone functionals ρ (see, Definition 9) with the Fatou property can be extended uniquely to a $\|.\|_1$-lower semicontinuous dilatation monotone functionals, which can preserves monotonicity, (quasi)convexity and cash-invariance of ρ. Our findings complement extension and continuity results for (quasi)convex law-invariant functional which is proven in [47,52,67].

2.2.3 Preliminaries and Framework

In this section, we recall a few essential terminologies which are used throughout our investigations in this Chapter.

Setting and Notation

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a fixed non-atomic probability space. By convention, without further notice, all equalities and inequalities are understood in the $\mathbb{P}$-almost sure (a.s.) sense. We denote by $L^1 := L^1(\Omega, \mathcal{F}, \mathbb{P})$ the space of integrable random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ equipped by $\|.\|_1$-norm. The space of bounded random variables is denoted by $L^\infty := L^\infty(\Omega, \mathcal{F}, \mathbb{P})$.

In this study, we assume the domain space $\mathcal{X}$ is a subset of L^1. The sole assumption we impose on the domain is that $\mathcal{X}$ contains the space of simple random variables $\mathcal{L}$, i.e.

$$\mathcal{L} \subset \mathcal{X} \subset L^1. \tag{2.1}$$

Definition 4. *A subspace $\mathcal{X}$ of L^1 is said to be an ideal of L^1 whenever $|X| \leq |Y|$ and $Y \in \mathcal{X}$ imply $X \in \mathcal{X}$ for all $X \in L^1$.*

All Orlicz space, Orlicz hearts space, as well as the L^p, $p \in [1, \infty]$ spaces are of this type.

Definition 5. *For any subset $\mathcal{S}$ of L^1, the smallest ideal of L^1 including $\mathcal{S}$ is called the ideal generated by $\mathcal{S}$, denote by $\mathcal{I}(\mathcal{S})$, is given by*

$$\mathcal{I}(\mathcal{S}) = \left\{ X \in L^1 : n \in \mathbb{N}, \textit{there exist } \; X_1, \ldots, X_n \in \mathcal{S} \textit{ and } \lambda_i \in \mathbb{R}_+ \textit{ such that } |X| \leq \sum_{i=1}^{n} \lambda_i |X_i| \right\}.$$

Note that by definition $\mathcal{S} \subseteq \mathcal{I}(\mathcal{S})$. The existence of such a smallest ideal follows from the fact that L^1 itself is an ideal including $\mathcal{S}$, and because the intersection of a family of ideals is an ideal again.

In the following, we recall some fundamental notions of functional ρ. Note that there are several possible properties. We only focus on those that are used in this thesis.

Definition 6. *Let $\rho : \mathcal{X} \to [-\infty, \infty]$ be a (risk) functional then ρ may fulfill the following properties:*

(A0) Proper: $\rho(X) > -\infty$ for every $X \in \mathcal{X}$, and $\rho(X) < \infty$ for some $X \in \mathcal{X}$.

(A1) Monotonicity: For any $X, Y \in \mathcal{X}$ with $X \geq Y$, $\rho(X) \leq \rho(Y)$.

(A2) Cash-invariance: For any $X \in \mathcal{X}$ and $c \in \mathbb{R}$ with $X + c \in \mathcal{X}$

$$\rho(X + c) = \rho(X) - c.$$

(A3) Convexity: For any $\lambda \in (0,1)$ such that $\lambda X + (1-\lambda)Y \in \mathcal{X}$,

$$\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y), \quad \textit{for any } X, Y \in \mathcal{X}.$$

(A4) Quasiconvex: For any $\lambda \in (0,1)$ such that $\lambda X + (1-\lambda)Y \in \mathcal{X}$,

$$\rho(\lambda X + (1-\lambda)Y) \leq \max\{\rho(X), \rho(Y)\}, \quad \textit{for any } X, Y \in \mathcal{X}.$$

Finite Conditional Expectation

Finite conditional expectation plays an essential role in this study. For the convenience of the reader, we recall here the definition of finite conditional expectation.

Definition 7. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a measurable space. A finite collection $\pi = \{A_1, \ldots, A_n\}$ of nonempty subsets of Ω is said to be a finite measurable partition of Ω if*

(i) $A_i \in \mathcal{F}$, for any i,

(ii) $A_i \cap A_j = \emptyset$, for all $1 \leq i < j$,

(iii) $\bigcup_{i=1}^{n} A_i = \Omega$.

The finite σ-algebra consisting of all possible unions of the A_i (including the empty set) is called the finite σ-algebra generated by π and is denoted by $\sigma(\pi) := \sigma(\{A_1, \ldots, A_n\})$.

We consider a finite measurable partition of Ω whose members have non-zero probabilities. We write Π for the collection of all such π.

Definition 8. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Suppose that $\sigma(\pi)$ is a finite σ- algebra generated by partition $\pi = \{A_1, A_2, \ldots, A_n\}$. Then conditional expectation of $X \in L^1$ with respect to $\sigma(\pi)$ is the simple random variable and defines as follows*

$$\mathbb{E}[X|\sigma(\pi)](\omega) := \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}(\omega), \quad \textit{for any } X \in L^1, \tag{2.2}$$

where $a_i = \mathbb{E}(X|A_i) = \frac{\mathbb{E}(X\mathbb{1}_{A_i})}{\mathbb{P}(A_i)}$.

Note that $\mathbb{E}[X|\sigma(\pi)]$ is well-defined and a simple function and thus it is bounded. For convenience, we write

$$\mathbb{E}[X|\pi] := \mathbb{E}[X|\sigma(\pi)].$$

2.3 Dilatation Monotone Maps

A crucial step in the proof of the results in this chapter is that the underlying assumptions force the risk measure to be dilatation monotone. This notation of risk aversion, which is a sort of (convex) ordering preference of decision-makers over risks, was introduced in [21, 50].

Recall that for any $X \in \mathcal{X}$ and any finite σ-algebra $\pi \in \Pi$, $\mathbb{E}[X|\pi]$ is a simple function, so the conditional expectation $\mathbb{E}[X|\pi] \in \mathcal{L} \subset \mathcal{X}$. Therefore one can define the dilatation monotone map ρ as the following.

Definition 9. *A map $\rho : \mathcal{X} \to [-\infty, \infty]$ is called (finite) dilatation monotone if*

$$\rho(\mathbb{E}[X|\pi]) \leq \rho(X), \quad \textit{for all } X \in \mathcal{X}, \quad \textit{and for all } \ \pi \in \Pi.$$

From a financial point of view, this definition is very natural. Since we have more available information by π for the position $\mathbb{E}[X|\pi]$ (more determined) than position X by itself, therefore position $\mathbb{E}[X|\pi]$ should hold less risk than position X.

2.3.1 A Connection Between Dilatation Monotonicity and Other Risk Aversions

This section gathers an interesting result that links dilatation monotonicity to the other risk aversion, more precisely convex order and law invariant; for more detail and a connection to other risk aversions, we refer the reader to [21, 52, 55, 64].

The preference relations between random variables can be established using the convex order relation, which is successfully applied in many areas due to its consistency with risk-averse preferences.

Definition 10. *Let $X, Y \in L^1$, X is said to be dominates Y in the convex order, written as $X \succeq_{cx} Y$, if*

$$\mathbb{E}[X] = \mathbb{E}[Y] \quad \textit{and} \quad \mathbb{E}[(X-t)^+] \geq \mathbb{E}[(Y-t)^+], \quad \textit{for all } t \in \mathbb{R}.$$

Equivalently, $X \succeq_{cx} Y$ if and only if for all convex functions $\varphi : \mathbb{R} \to \mathbb{R}$,

$$E[\varphi(X)] \geq E[\varphi(Y)].$$

Note that, using Jensen's inequality for the conditional expectation, for all convex functions $\varphi : \mathbb{R} \to \mathbb{R}$, we have

$$\mathbb{E}[\varphi(\mathbb{E}[X|\pi])] \leq \mathbb{E}[\mathbb{E}[\varphi(X)|\pi]] = E[\varphi(X)], \quad \text{for any } X \in L^1 \text{ and } \pi \in \Pi. \tag{2.3}$$

Thus

$$X \succeq_{cx} \mathbb{E}[X|\pi], \quad \text{for any } X \in L^1 \text{ and } \pi \in \Pi. \tag{2.4}$$

Definition 11. *A map $\rho : \mathcal{X} \to [-\infty, \infty]$ is called $\succeq_{cx}$-monotone (preserve $\succeq_{cx}$) if $X \succeq_{cx} Y$ then*

$$\rho(X) \geq \rho(Y).$$

Dilatation monotonicity is closely related to the convex order monotonicity. Using (2.4), It is easy to check that if a map ρ preserves $\succeq_{cx}$, then it is a dilatation monotone map.

Recall that, two random variables X and Y on probability space $(\Omega, \mathcal{F}, \mathbb{P})$ are equally distributed, denoted by $X \sim Y$, if $\text{law}_X = \text{law}_Y$, where law_X denotes the distribution of a random variable X under $\mathbb{P}$.

Definition 12. *A map $\rho : \mathcal{X} \to [-\infty, \infty]$ is said to be law invariant if for any $X, Y \in \mathcal{X}$ such that X,Y have the same distribution under $\mathbb{P}$, then*

$$\rho(X) = \rho(Y).$$

Namely, for any $X \in \mathcal{X}$, $\rho(X)$ depends only on the distribution of X.

Law invariance property ensures that two positions X and Y that have the same probability function have equal risks ($\rho(X) = \rho(Y)$). This characteristic is essential for practical risk measurements.

An interesting relationship between dilatation monotone and law invariant maps has been proved when $\mathcal{X} = L^\infty$ in [21].

Theorem 13. *([21], Theorem 1.1) On any atomless probability space, any $\|.\|_\infty$-continuous dilatation monotone map $\rho : L^\infty \to \mathbb{R}$ is law invariant.*

This result then extends to $\mathcal{X} = L^p$. In [64], the authors show that for any lower semicontinuous function on the L^p spaces, the dilatation monotonicity is equivalent to the law invariant of that function (For more general spaces, we refer the reader to [10, 52]). For the convenience of the reader, we provide partial proof of this result in the following.

Theorem 14. *([64], Theorem 2.1) Let $\rho : L^p \to (-\infty, \infty]$ be a convex norm-lower semi continues functional where $p \in [1, \infty]$. Then the following are equivalent:*

(i) ρ *is law-invariant.*

(ii) ρ *is $\succeq_{cx}$-monotone.*

(iii) ρ *is dilatation monotone.*

Proof Sketch: $(i) \Rightarrow (ii)$ [22], Theorem 4.1. $(ii) \Rightarrow (iii)$ It is easy to check.
$(iii) \Rightarrow (i)$
Step 1: When $\mathcal{X} = L^\infty$, By Theorem 13 any dilatation monotone, norm continuous functional ρ is law invariant. Indeed, for every $X, Y \in L^\infty$ with $X \sim Y$ and $\epsilon > 0$, there exist a finite sequence $(\mathcal{H}_k)_{k=1}^{N_\epsilon}$ of sub σ-algebra of $\mathcal{F}$ such that

$$\|X - \mathbb{E}[\dots \mathbb{E}[\mathbb{E}[Y|\mathcal{H}_1]|\mathcal{H}_2] \dots |\mathcal{H}_{N\epsilon}]\|_\infty < \epsilon.$$

Then by passing to a subsequence $(\mathcal{H}_k^n)_{k=1}^{N_\epsilon(n)}$, we have

$$\mathbb{E}[\dots \mathbb{E}[\mathbb{E}[Y|\mathcal{H}_1^n]|\mathcal{H}_2^n] \dots |\mathcal{H}_{N_\epsilon(n)}^n], \quad n \in \mathbb{N},$$

converge to X. Combining dilatation monotone, and continuity of ρ, we get

$$\rho(X) \leq \liminf_n \rho(\mathbb{E}[\ldots \mathbb{E}[\mathbb{E}[Y|\mathcal{H}_1^n]|\mathcal{H}_2^n]\ldots|\mathcal{H}_{N_\epsilon(n)}^n]) \leq \rho(Y).$$

Similarly, $\rho(Y) \leq \rho(X)$. So ρ is law invariant for any $X, Y \in L^\infty$.

Step 2: When $X, Y \in L^p$, $p \in [1, \infty)$ with $X \sim Y$. Then there exist σ-algebra $\mathcal{A}_n$, $\mathcal{B}_n$, such that

$$\mathbb{E}[X|\mathcal{A}_n] \sim \mathbb{E}[Y|\mathcal{B}_n], \quad \text{for each } n \in \mathbb{N}, \tag{2.5}$$

where $(\mathbb{E}[X|\mathcal{A}_n])_{n\in\mathbb{N}}$, $(\mathbb{E}[Y|\mathcal{B}_n])_{n\in\mathbb{N}}$ are bounded and converge in L^p to X, Y respectively.

Since ρ is lower semicontinuity and dilatation monotone then

$$\rho(Y) = \lim_n \rho(\mathbb{E}[Y|\mathcal{B}_n]), \quad \rho(X) = \lim_n \rho(\mathbb{E}[X|\mathcal{A}_n]). \tag{2.6}$$

And finally, since $(\mathbb{E}[X|\mathcal{A}_n])_n, (\mathbb{E}[Y|\mathcal{B}_n])_n \in L^\infty$, and ρ is law invariant on L^∞, using (2.5) and (2.6) and the result from Step 1, we have

$$\rho(X) = \lim_n \rho(\mathbb{E}[X|\mathcal{A}_n]) = \lim_n \rho(\mathbb{E}[Y|\mathcal{B}_n]) = \rho(Y).$$

$\square$

2.3.2 An Approximation by a Finite Conditional Expectation

Now, we established one of our main results for this Chapter which is related to convergence of conditional expectations, and it is of fundamental importance for the study of dilatation monotone risk functionals. First, we recall a preliminary result that deals with the convergence of finite conditional expectation in the L^∞ space. When $\mathcal{X} = L^\infty$, it is common-known that any $X \in L^\infty$ can be approximated by a finite conditional expectation, which is recorded in the following lemma (see, [32, 43, 63]).

Lemma 15. *([32], Lemma 3.1) For any $X \in L^\infty$ and $\epsilon > 0$, there exists $\pi \in \Pi$ such that*

$$\|\mathbb{E}[X|\pi] - X\|_\infty < \epsilon.$$

Definition 16. *For any $\Omega' \in \mathcal{F}$ with $\mathbb{P}(\Omega') > 0$, $(\Omega', \mathcal{F}_{|\Omega'}, \mathbb{P}_{|\Omega'})$ is called the localized probability space, where $\mathcal{F}_{|\Omega'} := \{E \in \mathcal{F} : E \subset \Omega'\}$ and $\mathbb{P}_{|\Omega'} : \mathcal{F}_{|\Omega'} \to [0,1]$ is defined by*

$$\mathbb{P}_{|\Omega'}(E) := \mathbb{E}[E|\Omega'] = \frac{\mathbb{P}(E)}{\mathbb{P}(\Omega')}.$$

Note that $(\Omega', \mathcal{F}_{|\Omega'}, \mathbb{P}_{|\Omega'})$ is also non-atomic. For any X, we denote by $X_{|\Omega'}$ the random variable on $(\Omega', \mathcal{F}_{|\Omega'}, \mathbb{P}_{|\Omega'})$ obtained by restricting X to Ω'.

Now, we established one of our first main results, which can be of individual interest. An interesting aspect of this Lemma is that one can approximate every $X \in \mathcal{X} \subset L^1$ with a sequence of conditional expectations respect to finite σ-algebra in such a way that $\sup_{n\in\mathbb{N}} |\mathbb{E}[X|\pi_n]| \in \mathcal{I}(\{X, \mathbb{1}\})$.

Lemma 17. *For any $X \in L^1$, there exists $(\pi_n)_{n\in\mathbb{N}} \subset \Pi$ and $k \in \mathbb{R}_+$ such that*

$$\mathbb{E}[X|\pi_n] \xrightarrow{a.s.} X, \quad \textit{and} \quad |\mathbb{E}[X|\pi_n]| \leq |X| + k, \quad \textit{for all } n \in \mathbb{N}.$$

Proof. Let $X \in L^1$ and $k_1 \in \mathbb{R}_+$ be such that $\mathbb{P}(|X| \leq k_1) > \frac{1}{2}$. Pick any $0 < \epsilon < \frac{1}{2}$. Since $(\Omega, \mathcal{F}, \mathbb{P})$ is non-atomic, we can find $A \in \mathcal{F}$ such that

$$A \subset \{|X| \leq k_1\}, \quad \text{and} \quad \mathbb{P}(A) = \epsilon. \tag{2.7}$$

In addition, since $X \in L^1$, we can find $k_2 \in \mathbb{R}$ such that

$$k_2 > k_1, \quad \text{and} \quad \mathbb{E}[|X|\mathbb{1}_{|X|>k_2}] < \epsilon. \tag{2.8}$$

Put $\Omega' := \{|X| \leq k_2\} \setminus A$. Note that since $k_2 > k_1$, using (2.7), we have

$$\begin{aligned}\mathbb{P}(\Omega') &= \mathbb{P}(|X| \leq k_2) - \mathbb{P}(A) \\ &\geq \mathbb{P}(|X| \leq k_1) - \mathbb{P}(A) > 0.\end{aligned}$$

As $\mathbb{P}(\Omega') > 0$, by Definition 16, we can consider the localized probability space $(\Omega', \mathcal{F}_{|\Omega'}, \mathbb{P}_{|\Omega'})$. For the localized probability space $(\Omega', \mathcal{F}_{|\Omega'}, \mathbb{P}_{|\Omega'})$, we have that $X\mathbb{1}_{\Omega'} \in L^\infty|_{|\Omega'} := L^\infty(\Omega', \mathcal{F}_{|\Omega'}, \mathbb{P}_{\Omega'})$. By applying Lemma 15, since $X\mathbb{1}_{\Omega'} \in L^\infty|_{\Omega'}$, we can find a measurable partition $\pi' = \{B_1, B_2, ..., B_n\}$ of Ω' such that $\mathbb{P}(B_i) > 0$ for each $i = 1, \ldots, n$ and

$$\left| \sum_{B \in \pi'} \frac{\mathbb{E}[X\mathbb{1}_B]}{\mathbb{P}(B)} \mathbb{1}_B - X\mathbb{1}_{\Omega'} \right| < \epsilon \mathbb{1}_{\Omega'}. \tag{2.9}$$

Note that $\Omega \setminus \Omega' = \{|X| > k_2\} \cup A$, so we have

$$\mathbb{P}(\Omega \setminus \Omega') \geq \mathbb{P}(A) = \epsilon > 0.$$

Thus

$$\pi = \{B_1, B_2, \ldots, B_n, \Omega \setminus \Omega'\} \in \Pi.$$

By applying (2.9), we obtain

$$\begin{aligned}|\mathbb{E}[X|\pi]\mathbb{1}_{\Omega'}| &\leq |\mathbb{E}[X|\pi]\mathbb{1}_{\Omega'} - X\mathbb{1}_{\Omega'}| + |X\mathbb{1}_{\Omega'}| \\ &= \left| \sum_{B \in \pi'} \frac{\mathbb{E}[X\mathbb{1}_B]}{\mathbb{P}(B)} \mathbb{1}_B - X\mathbb{1}_{\Omega'} \right| + |X\mathbb{1}_{\Omega'}| \\ &< \epsilon \mathbb{1}_{\Omega'} + |X| \qquad (2.10) \\ &< \frac{1}{2}\mathbb{1}_{\Omega'} + |X|.\end{aligned}$$

Now, we will find an upper bound for $|\mathbb{E}[X|\pi]\mathbb{1}_{\Omega\setminus\Omega'}|$. By using (2.7) and (2.8), we have

$$\begin{aligned}
\mathbb{E}[|X|\mathbb{1}_{\Omega\setminus\Omega'}] &= \mathbb{E}[|X|\mathbb{1}_{A\cup\{|X|>k_2\}}] \\
&= \mathbb{E}[|X|\mathbb{1}_{A}] + \mathbb{E}[|X|\mathbb{1}_{\{|X|>k_2\}}] \\
&< \epsilon(k_1+1).
\end{aligned} \tag{2.11}$$

In view of $\mathbb{P}(\Omega\setminus\Omega') \geq \epsilon$, and by applying (2.11), we get

$$\begin{aligned}
|\mathbb{E}[X|\pi]\mathbb{1}_{\Omega\setminus\Omega'}| &\leq \frac{\mathbb{E}[|X|\mathbb{1}_{\Omega\setminus\Omega'}]}{\mathbb{P}(\Omega\setminus\Omega')}\mathbb{1}_{\Omega\setminus\Omega'} \\
&< (k_1+1)\mathbb{1}_{\Omega\setminus\Omega'}.
\end{aligned}$$

Therefore, by invoking (2.10), it follows that

$$\begin{aligned}
|\mathbb{E}[X|\pi]| &\leq |\mathbb{E}[X|\pi]\mathbb{1}_{\Omega'}| + |\mathbb{E}[X|\pi]\mathbb{1}_{\Omega\setminus\Omega'}| \\
&\leq |X| + k_1 + 1.
\end{aligned} \tag{2.12}$$

Now we claim

$$\|\mathbb{E}[X|\pi] - X\|_1 < \epsilon(3+2k_1). \tag{2.13}$$

Indeed, by applying (2.9) and (2.11), we get the following

$$\begin{aligned}
\|\mathbb{E}[X|\pi] - X\|_1 &= \|(\mathbb{E}[X|\pi] - X)\mathbb{1}_{\Omega'}\|_1 + \left\|(\mathbb{E}[X|\pi] - X)\mathbb{1}_{\Omega\setminus\Omega'}\right\|_1 \\
&\leq \epsilon + \left\|\mathbb{E}[X|\pi]\mathbb{1}_{\Omega\setminus\Omega'}\right\|_1 + \left\|X\mathbb{1}_{\Omega\setminus\Omega'}\right\|_1 \\
&= \epsilon + \left\|\mathbb{E}[X\mathbb{1}_{\Omega\setminus\Omega'}|\pi]\right\|_1 + \left\|X\mathbb{1}_{\Omega\setminus\Omega'}\right\|_1 \\
&\leq \epsilon + 2\left\|X\mathbb{1}_{\Omega\setminus\Omega'}\right\|_1 \\
&< \epsilon + 2\epsilon(k_1+1) \\
&= \epsilon(3+2k_1).
\end{aligned}$$

Now letting, e.g., $\epsilon = \frac{1}{n}$ in (2.12) and (2.13), we obtain $(\pi_n)_{n\in\mathbb{N}} \subset \Pi$ such that

$$\sup_{n\in\mathbb{N}} |\mathbb{E}[X|\pi_n| \leq |X| + (k_1+1)\mathbb{1} \quad \text{and} \quad \mathbb{E}[X|\pi_n] \xrightarrow{\|\cdot\|_1} X.$$

Since $\mathbb{E}[X|\pi_n] \xrightarrow{\|\cdot\|_1} X$, by passing to a subsequence $(\pi_{n_k})_{k\in\mathbb{N}}$, we have

$$\mathbb{E}[X|\pi_{n_k}] \xrightarrow{a.s} X, \quad \text{and} \quad |\mathbb{E}[X|\pi_{n_k}]| \leq |X| + k.$$

By replacing $(\pi_n)_{n\in\mathbb{N}}$ with $(\pi_{n_k})_{k\in\mathbb{N}}$, we get the desire result. $\square$

2.3.3 On Continuity of Dilatation Monotone Maps

Continuity plays an essential role in risk measure theory. From a practical perspective, continuity is necessary since if the risk measureρ fails to be continuous at some position X, then a slight change or misstatement of X might lead to a dramatically different capital requirement. From a mathematical point of view, continuity is a useful property in the context of dual representations, which is very important in optimization problems. One of the technical conditions on the continuity of risk measure is the Fatou property. In this section, we investigate a relationship between the Fatou property and τ-lower semicontinuity ($\|\|_1$-lower semicontinuity, $\sigma(L^1, \mathcal{L})$-lower semicontinuity) of a dilatation monotone functional ρ base on our model assumption[2].

Let τ be a topology on $\mathcal{X}$. We say that ρ is τ-lower semicontinuous whenever the sublevel set

$$\{\rho \leq \lambda\} := \{X \in \mathcal{X} : \rho(X) \leq \lambda\} \tag{2.14}$$

is τ-closed for each $\lambda \in \mathbb{R}$.

We recall that if τ is a metrizable topology (see, for example, [16], Lemma 2.42), then ρ is τ-lower

[2]On early results connecting lower semicontinuity with the Fatou property, we suggest the reader to [12].

semicontinuous if and only if, for any sequence $(X_n)_{n\in\mathbb{N}}$ in $\mathcal{X}$ and $X \in \mathcal{X}$, we have

$$X_n \xrightarrow{\tau} X,$$

implies

$$\rho(X) \leq \liminf_n \rho(X_n)$$

In the following, we define the Fatou property which will be vitally important for our investigations in this Chapter.

Definition 18. *A functional $\rho : \mathcal{X} \to [-\infty, \infty]$ has the Fatou property if for any sequence $(X_n)_{n\in\mathbb{N}}$ in $\mathcal{X}$ and $X \in \mathcal{X}$ we have*

$$X_n \xrightarrow{a.s.} X, \quad \text{and} \quad \sup_{n\in\mathbb{N}} |X_n| \in \mathcal{I}(\mathcal{X}). \tag{2.15}$$

Then implies

$$\rho(X) \leq \liminf_n \rho(X_n). \tag{2.16}$$

Remark 19. *We note here that in the standard framework where $\mathcal{X}$ is a rearrangement invariant space ($X \in \mathcal{X}$ and $X \sim Y$ then $Y \in \mathcal{X}$), or more generally, an ideal of L^1, the supremum taken in (2.15) is an element of $\mathcal{X}$, and thus, the definition of the Fatou property we give here coincides with the one used in the literature (see, e.g. [34] and the references therein).*

Note that any functional with the Fatou property is automatically $\|.\|_p$-lower semicontinuous. This holds because every norm convergent sequence has a dominated subsequence that converges almost surely (see, for example, [16], Theorem 13.6). In [43], Jouini et al. showed that for any proper convex law-invariant functional $\rho : L^\infty \to (-\infty, \infty]$, norm lower semicontinuity of ρ implies the Fatou property. In general, for $\mathcal{X} = L^p,\ 1 \leq p \leq \infty$, the Fatou property is equivalent to $\|.\|_p$-lower semicontinuity, for any proper convex risk measure ([44], Theorem 3.3).

In the following proposition, we show that if ρ is dilatation monotone then the Fatou property is, in fact, equivalent to lower semicontinuity with respect to the relative $\sigma(L^1, \mathcal{L})$ topology on $\mathcal{X}$.

Proposition 20. *Let $\rho : \mathcal{X} \to [-\infty, \infty]$ be dilatation monotone. Then the following are equivalent.*

(i) ρ *has the Fatou property.*

(ii) ρ *is lower semicontinuous with respect to the relative* $\sigma(L^1, \mathcal{L})$ *topology on* $\mathcal{X}$.

(iii) ρ *is lower semicontinuous with respect to the relative* $||\cdot||_1$ *topology on* $\mathcal{X}$.

If any of the above holds, then for any $X \in \mathcal{X}$,

$$\rho(X) = \sup_{\pi \in \Pi} \rho(\mathbb{E}[X|\pi]), \tag{2.17}$$

and for any $(\pi_n)_{n\in\mathbb{N}} \subset \Pi$,

$$\mathbb{E}[X|\pi_n] \xrightarrow{\sigma(L^1,\mathcal{L})} X \implies \rho(X) = \lim_n \rho(\mathbb{E}[X|\pi_n]). \tag{2.18}$$

Proof. $(i) \Rightarrow (ii)$. Let assume ρ has the Fatou property and $\mathcal{A} := \{X \in \mathcal{X} : \ \rho(X) \le \lambda\}$ be a sublevel set. Let (X_α) be any net in $\mathcal{A}$ converging to some $X \in \mathcal{X}$ in the $\sigma(L^1, \mathcal{L})$ -topological sense, i.e.,

$$X_\alpha \xrightarrow{\sigma(L^1,\mathcal{L})} X \in \mathcal{X}.$$

We will show $X \in \mathcal{A}$.

Since $X \in \mathcal{X} \subset L^1$, by Lemma 17, we can find $k \in \mathbb{R}_+$ and $(\pi_m)_{m\in\mathbb{N}} \subset \Pi$ such that

$$\mathbb{E}[X|\pi_m] \xrightarrow{a.s.} X, \quad \text{and} \quad \sup_{m\in\mathbb{N}} |\mathbb{E}[X|\pi_m]| \le |X| + k.$$

By assumption, ρ has the Fatou property. Thus by using (2.15), we have

$$\rho(X) \le \liminf_m \rho(\mathbb{E}[X|\pi_m]). \tag{2.19}$$

Now, we claim

$$\mathbb{E}[X_\alpha|\pi] \xrightarrow{\|\cdot\|_1} \mathbb{E}[X|\pi], \quad \text{for any } \pi \in \Pi. \tag{2.20}$$

To verify this, note that for any $\pi := \{A_1, \ldots, A_n\} \in \Pi$, we have

$$\begin{aligned}
\|\mathbb{E}[X_\alpha|\pi] - \mathbb{E}[X|\pi]\|_1 &= \|\mathbb{E}[X - X_\alpha|\pi]\|_1 \\
&= \left\| \sum_{A\in\pi} \frac{\mathbb{E}[X_\alpha \mathbb{1}_A - X\mathbb{1}_A]}{\mathbb{P}(A)} \mathbb{1}_A \right\|_1 \\
&\leq \mathbb{E}\left[\sum_{A\in\pi} \left| \frac{\mathbb{E}\left[X_\alpha \mathbb{1}_A - X\mathbb{1}_A\right]}{\mathbb{P}(A)} \right| \mathbb{1}_A \right] \\
&= \sum_{A\in\pi} \mathbb{E}\left[\frac{\left|\mathbb{E}\left[X_\alpha \mathbb{1}_A - X\mathbb{1}_A\right]\right|}{\mathbb{P}(A)} \mathbb{1}_A \right] \\
&= \sum_{A\in\pi} |\mathbb{E}[X_\alpha \mathbb{1}_A - X\mathbb{1}_A]| \to 0.
\end{aligned}$$

Now by using (2.20), we can choose $(\alpha_n)_{n\in\mathbb{N}}$ such that

$$\|\mathbb{E}[X_{\alpha_n}|\pi] - \mathbb{E}[X|\pi]\|_1 \leq \frac{1}{2^n}, \quad \text{for all } n \in \mathbb{N}.$$

Then

$$Y := \sum_{k=1}^{\infty} |\mathbb{E}[X_{\alpha_k} - X|\pi]| \in L^1. \tag{2.21}$$

Now let $\mathcal{R}_\pi$ be the space of random variables measurable with respect to $\sigma(\pi)$, then

$$Y_n := \sum_{k=1}^{n} |\mathbb{E}[X_{\alpha_k} - X|\pi]| \in \mathcal{R}_\pi, \quad \text{for any } n \in \mathbb{N}.$$

Since $\mathcal{R}_\pi$ is a finite dimensional subspace of L^1 , it follows that $\mathcal{R}_\pi$ is a closed subspace of L^1 and thus $Y \in \mathcal{R}_\pi$. Moreover,

$$|\mathbb{E}[X_{\alpha_n}|\pi] - \mathbb{E}[X|\pi]| \leq \sum_{k=n}^{\infty} |\mathbb{E}[X_{\alpha_k}|\pi] - \mathbb{E}[X|\pi]| \xrightarrow{a.s.} 0,$$

and we have

$$\sup_{n\in\mathbb{N}}\{|\mathbb{E}[X_{\alpha_n}|\pi]|\} \leq Y + |\mathbb{E}[X|\pi]| \in \mathcal{L} \subset \mathcal{X} \subset \mathcal{I}(\mathcal{X}).$$

Therefore, by applying (2.15) to π_m again, we get

$$\rho(\mathbb{E}[X|\pi_m]) \leq \liminf_n \rho(\mathbb{E}[X_{\alpha_n}|\pi_m]), \quad \text{for any } m \in \mathbb{N}.$$

Applying dilatation monotonicity of ρ to the right hand side, we obtain

$$\rho(\mathbb{E}[X|\pi_m]) \leq \liminf_n \rho(X_{\alpha_n}), \quad \text{for any } m \in \mathbb{N}.$$

Finally, by using (2.19), we get

$$\rho(X) \leq \liminf_n \rho(X_{\alpha_n}) \leq \lambda.$$

Thus $X \in \mathcal{A} = \{X \in \mathcal{X} : \rho(X) \leq \lambda\}$.

$(ii) \Rightarrow (iii)$. This is immediate since the $\sigma(L^1, \mathcal{L})$ topology is weaker than the $\|\cdot\|_1$-topology.

$(iii) \Rightarrow (i)$. Let $(X_n)_{n\in\mathbb{N}} \subset \mathcal{X}$ and $X \in \mathcal{X}$ such that

$$X_n \xrightarrow{a.s.} X, \quad \text{and} \quad \sup_{n\in\mathbb{N}} |X_n| \in \mathcal{I}(\mathcal{X}).$$

By the Dominated Convergence Theorem, we have

$$X_n \xrightarrow{\|\cdot\|_1} X. \tag{2.22}$$

By assumption, ρ is lower semicontinuous with respect to the relative $||\cdot||_1$-topology on $\mathcal{X}$, so we have

$$\rho(X) \leq \liminf_n \rho(X_n).$$

Therefore, ρ has the Fatou property.

To verify (2.18), let

$$\mathbb{E}[X|\pi_n] \xrightarrow{\sigma(L^1,\mathcal{L})} X$$

and note that by (ii) and the dilatation monotonicity of ρ, we have that

$$\rho(X) \leq \liminf_n \rho(\mathbb{E}[X|\pi_n]) \leq \limsup_n \rho(\mathbb{E}[X|\pi_n]) \leq \rho(X).$$

In particular, it follows that

$$\rho(X) = \lim_n \rho(\mathbb{E}[X|\pi_n]).$$

Finally, (2.17) follows by the dilatation mononicity and the Fatou property of ρ, and Lemma 17. $\square$

2.4 On Extension of Dilatation Monotone Maps

In this section, we provide the key extension result for functionals $\rho : \mathcal{X} \to (-\infty, \infty]$. Recall that, in celebrated work, Filipović, and Svindland in [29] propound an idea of extension of functional ρ defines on the L^∞ space to a functional on the L^1 space. While many extension results, such as [10, 17, 29, 45], required the functional ρ be a (quasi)convex function, the novelty of our result is that we do not impose any (quasi)convexity assumptions on the underlying risk functional ρ.

2.4.1 On Extension Theory

In order to enlarge the domain of the map $\rho : \mathcal{X} \to (-\infty, \infty]$ to the L^1 space, we first construct a candidate function. The equation (2.17) motivates us to define the extension functional associated with ρ.

Definition 21. *For any functional $\rho : \mathcal{X} \to (-\infty, \infty]$, the extension functional of ρ to L^1 defines as the following*

$$\overline{\rho}(X) := \sup_{\pi \in \Pi} \rho(\mathbb{E}[X|\pi]), \quad \text{for all } X \in L^1. \tag{2.23}$$

The following theorem provides an extension result for the dilatation monotone functional ρ. In particular, the following theorem shows that every dilatation monotone functional ρ on $\mathcal{X}$ with the Fatou property can uniquely extend to the entire space of L^1 without losing its properties.

Theorem 22. *Let $\rho : \mathcal{X} \to (-\infty, \infty]$ be a dilatation monotone functional with the Fatou property and $r\bar{h}o$ as in (2.23). Then the extension functional $\overline{\rho} : L^1 \to (-\infty, \infty]$ has the following properties.*

(i) $\overline{\rho}$ is dilatation monotone.

(ii) $\overline{\rho}$ is $\sigma(L^1, \mathcal{L})$ lower semicontinuous, and $\overline{\rho}$ is a unique extension of ρ to L^1

(iii) $\overline{\rho}$ is law-invariant

Moreover, If ρ is (quasi)convex and monotone, and cash-invariant then

(iv) $\bar{\rho}$ preserves (quasi)convexity, monotonicity and cash-invariant of ρ.

Proof. First, we show that $\overline{\rho} : L^1 \to (-\infty, \infty]$. For any $X \in L^1$, we have

$$\mathbb{E}[X|\pi] \in \mathcal{L} \subset \mathcal{X}, \quad \text{for any } \pi \in \Pi.$$

Note that $\rho > -\infty$ on $\mathcal{X}$, since $\mathbb{E}[X|\pi] \in \mathcal{X}$ so $\rho(\mathbb{E}[X|\pi]) > -\infty$, for all $X \in L^1$. By (2.23),

$$\overline{\rho}(X) = \sup_{\pi \in \Pi}(\mathbb{E}[X|\pi]) > -\infty, \quad \text{for any } X \in L^1.$$

Thus $\overline{\rho}$ is a map from L^1 to $(-\infty, \infty]$.

(i) We have to show that

$$\overline{\rho}(\mathbb{E}[X|\pi]) \leq \overline{\rho}(X), \quad \text{for any } X \in L^1, \ \pi \in \Pi.$$

First, we show that the extension functional $\overline{\rho}$ agrees with ρ on $\mathcal{X}$. Since ρ has the Fatou property, using equation (2.23) and (2.17), it is immediate to see that

$$\overline{\rho}(X) = \sup_{\pi \in \Pi} \rho(\mathbb{E}[X|\pi]) = \rho(X), \quad \text{for any } X \in \mathcal{X}. \tag{2.24}$$

Therefore, $\overline{\rho}|_{\mathcal{X}} = \rho$.

For any $X \in L^1$ and $\pi \in \Pi$, since $\mathbb{E}[X|\pi] \in \mathcal{L} \subset \mathcal{X}$, By using (2.23) and (2.24), we get

$$\overline{\rho}(\mathbb{E}[X|\pi]) = \rho(\mathbb{E}[X|\pi]) \leq \sup_{\pi \in \Pi} \rho(\mathbb{E}[X|\pi]) = \overline{\rho}(X).$$

This establishes dilatation monotonicity of $\overline{\rho}$ on L^1.

(ii) Enough to show that $\overline{\rho}$ is $|| \cdot ||_1$-lower semicontinuous, and thus by Proposition 20, we get that $\overline{\rho}$ is

$\sigma(L^1, \mathcal{L})$-lower semicontinuous. Indeed, let $X_n \xrightarrow{||\cdot||_1} X$ and fix some $\pi \in \Pi$. Then

$$\mathbb{E}[X_n|\pi] = \sum_{A\in\pi} \frac{\mathbb{E}[X_n \mathbb{1}_A]}{\mathbb{P}(A)} \mathbb{1}_A \xrightarrow{||\cdot||_1} \sum_{A\in\pi} \frac{\mathbb{E}[X \mathbb{1}_A]}{\mathbb{P}(A)} \mathbb{1}_A = \mathbb{E}[X|\pi].$$

Therefore, by applying the Fatou property of ρ for $(\mathbb{E}[X_n|\pi])_{n\in\mathbb{N}}$ and (2.23), we have

$$\begin{aligned}
\rho(\mathbb{E}[X|\pi]) &\leq \liminf_n \rho(\mathbb{E}[X_n|\pi]) \\
&\leq \liminf_n \sup_{\pi\in\Pi} \rho(\mathbb{E}[X_n|\pi]) \\
&\leq \liminf_n \overline{\rho}(X_n).
\end{aligned}$$

Taking supremum over π, we get

$$\overline{\rho}(X) = \sup_{\pi\in\Pi} \rho(\mathbb{E}[X|\pi]) \leq \liminf_n \overline{\rho}(X_n).$$

and thus $\overline{\rho}$ is $||\cdot||_1$-lower semicontinuous. Using Proposition 20, we obtain the desired result. The uniqueness of $\bar{\rho}$ is immediate by (2.17).

(iii) The law-invariance of $\overline{\rho}$ follows by [52], Theorem 18(ii).

(iv) It remains to prove that if ρ is (a) (quasi)convex, (b) monotone, and (c) cash-invariant, then $\overline{\rho}$ has the same properties.

(a) Assume that ρ is quasiconvex. Pick any $X_1, X_2 \in L^1$ and fix some $\lambda \in (0,1)$. Then

$$\begin{aligned}
\overline{\rho}(\lambda X_1 + (1-\lambda)X_2) &= \sup_{\pi\in\Pi} \rho(\mathbb{E}[\lambda X_1 + (1-\lambda)X_2|\pi]) \\
&= \sup_{\pi\in\Pi} \rho(\lambda\mathbb{E}[X_1|\pi] + (1-\lambda)\mathbb{E}[X_2|\pi]) \\
&\leq \sup_{\pi\in\Pi} \max\{\rho(\mathbb{E}[X_1|\pi]), \rho(\mathbb{E}[X_2|\pi)\} \\
&\leq \max\{\overline{\rho}(X_1), \overline{\rho}(X_2)\}.
\end{aligned}$$

Hence, $\overline{\rho}$ is quasiconvex as well. In a similar manner, one can show that ρ preserves convexity of ρ. In fact, for any $X, Y \in L^1$, and $\lambda \in (0,1)$, we have

$$\begin{aligned}
\overline{\rho}(\lambda X_1 + (1-\lambda)X_2) &= \sup_{\pi\in\Pi} \rho(\mathbb{E}[\lambda X_1 + (1-\lambda)X_2|\pi]) \\
&\leq \lambda \sup_{\pi\in\Pi} \rho(\mathbb{E}[X_1|\pi]) + (1-\lambda) \sup_{\pi\in\Pi} \mathbb{E}[X_2|\pi]) \\
&= \lambda\overline{\rho}(X_1) + (1-\lambda)\overline{\rho}(X_2).
\end{aligned}$$

(b) Pick $X, Y \in L^1$ such that $X \geq Y$. Since ρ is monotone, $\rho(X) \leq \rho(Y)$, then we obtain

$$\begin{aligned}
\overline{\rho}(X) &= \sup_{\pi\in\Pi} \rho(\mathbb{E}[X|\pi]) \\
&\leq \sup_{\pi\in\Pi} \rho(\mathbb{E}[Y|\pi]) \\
&= \overline{\rho}(Y).
\end{aligned}$$

(c) Pick $X \in L^1$, $c \in \mathbb{R}$. Let assume ρ is cash-invariant, $\rho(X+c) = \rho(X) - c$, then

$$\begin{aligned}
\overline{\rho}(X+c) &= \sup_{\pi\in\Pi} \rho(\mathbb{E}[X+c|\pi]) \\
&= \sup_{\pi\in\Pi} \rho(\mathbb{E}[X|\pi]) - c \\
&= \overline{\rho}(X) - c.
\end{aligned}$$

$\square$

In the following, we derive a dual representation for map $\rho : \mathcal{X} \to [-\infty, \infty]$ in terms of the conjugate function $\rho^{\#} : L^1 \to [-\infty, \infty]$ which is defined as the following

$$\rho^{\#}(Y) := \sup_{X \in \mathcal{L}} \{\mathbb{E}[XY] - \rho(X)\}, \qquad Y \in L^1. \tag{2.25}$$

Definition 23. *A subset $\mathcal{C}$ of L^1 is a dilatation monotone set if*

(i) $\mathcal{C}$ is non-empty set,

(ii) for all $X \in \mathcal{C}$ and $\pi \in \Pi$, $\mathbb{E}[X|\pi] \in \mathcal{C}$.

For example, if $\mathcal{C}$ convex, $\|.\|_1$-closed, law-invariant subset L^1, then $\mathcal{C}$ is dilation monotone set of L^1 ([43], Lemma 4.2).

It is easy to show that the conjugate functional $\rho^{\#}$ of a dilatation monotone functional ρ is a $\|.\|_1$-lower semicontinuous dilatation monotone functional.

Lemma 24. *Let $\rho : \mathcal{X} \to [-\infty, \infty]$ be dilatation monotone and $\rho^{\#}$ be as in (2.25), then*

(i) $\rho^{\#}$ is dilatation monotone.

(ii) $\rho^{\#}$ is $\|\cdot\|_1$-lower semicontinuous.

(iii) Let $\mathcal{C} \subset \mathcal{X}$ be a dilatation monotone subset of L^1. If ρ has the Fatou property, then for any $Y \in L^1$ such that

$$\mathbb{E}[|YX|] < \infty, \quad \textit{for all } X \in \mathcal{C},$$

we have

$$\sup_{X \in \mathcal{L} \cap \mathcal{C}} \{\mathbb{E}[XY] - \rho(X)\} = \sup_{X \in \mathcal{C}} \{\mathbb{E}[XY] - \rho(X)\}.$$

Proof. (i) Fix any $Y \in L^1$ and $\pi \in \Pi$. Note that $\mathbb{E}[Y|\pi] \in \mathcal{L} \subset L^1$. By and self disjoint property of conditional expectation, dilatation monotonicity of ρ and using (2.25), we have

$$\begin{aligned}
\rho^{\#}(\mathbb{E}[Y|\pi]) &= \sup_{X\in\mathcal{L}}\{\mathbb{E}[X\mathbb{E}[Y|\pi]] - \rho(X)\} \\
&= \sup_{X\in\mathcal{L}}\{\mathbb{E}[Y\mathbb{E}[X|\pi]] - \rho(X)\} \\
&\leq \sup_{X\in\mathcal{L}}\{\mathbb{E}[Y\mathbb{E}[X|\pi]] - \rho(\mathbb{E}[X|\pi])\} \\
&\leq \sup_{Z\in\mathcal{L}}\{\mathbb{E}[YZ] - \rho(Z)\} \\
&= \rho^{\#}(Y).
\end{aligned}$$

Thus $\rho^{\#}$ is dilatation monotone on L^1.

(ii) We show $\rho^{\#}$ is lower semicontinuous. Let $(Y_n)_{n\in\mathbb{N}} \subset L^1$ and $Y \in L^1$ such that

$$Y_n \xrightarrow{||\cdot||_1} Y.$$

Then

$$\lim_n(\mathbb{E}[XY_n] - \rho(X)) = \mathbb{E}[XY] - \rho(X), \quad \text{for all } X \in \mathcal{L}. \tag{2.26}$$

By using (2.25), we obtain

$$\rho^{\#}(Y_n) = \sup_{X\in\mathcal{L}}\{\mathbb{E}[XY_n] - \rho(X)\} \geq \mathbb{E}[XY_n] - \rho(X), \quad \text{for all } X \in \mathcal{L},$$

It follows that

$$\liminf_n \rho^{\#}(Y_n) \geq \mathbb{E}[XY] - \rho(X), \quad \text{for all } X \in \mathcal{L},$$

which implies

$$\liminf_n \rho^{\#}(Y_n) \geq \sup_{X\in\mathcal{L}}\{\mathbb{E}[XY] - \rho(X)\} = \rho^{\#}(Y).$$

This proves that $\rho^{\#}$ is $\|\cdot\|_1$-lower semicontinuous.

(iii) Fix $Y \in L^1$ such that

$$\mathbb{E}[|YX|] < \infty, \quad \text{for all } X \in \mathcal{C}. \tag{2.27}$$

It is clear that

$$\sup_{X \in \mathcal{L} \cap \mathcal{C}} \{\mathbb{E}[XY] - \rho(X)\} \leq \sup_{X \in \mathcal{C}} \{\mathbb{E}[XY] - \rho(X)\}. \tag{2.28}$$

For simplicity let

$$c := \sup_{X \in \mathcal{L} \cap \mathcal{C}} \{\mathbb{E}[XY] - \rho(X)\}.$$

Now pick any $X \in \mathcal{C}$. By Lemma 17, there exists a sequence $(\pi_n)_{n \in \mathbb{N}} \subset \Pi$ such that

$$\mathbb{E}[X|\pi_n] \xrightarrow{a.s.} X, \quad \text{and} \quad X^* := \sup_{n \in \mathbb{N}} \{|\mathbb{E}[X|\pi_n]|\} \in \mathcal{I}(\mathcal{C} \cup \{\mathbb{1}\}).$$

Therefore, by assumption (2.27), $\mathbb{E}[X^*|Y|] < \infty$. Then by the Dominated Convergence Theorem,

$$\mathbb{E}[YX] = \lim_n \mathbb{E}[Y\mathbb{E}[X|\pi_n]]. \tag{2.29}$$

Note that $\mathcal{C}$ is dilatation monotone set, so if $X \in \mathcal{C}$, then $\mathbb{E}[X|\pi] \in \mathcal{C}$. Therefore, $\mathbb{E}[X|\pi] \in \mathcal{C} \cap \mathcal{L}$ and we get

$$c \geq \mathbb{E}[\mathbb{E}[X|\pi_n]Y] - \rho(\mathbb{E}[X|\pi_n]), \quad \text{for any } n \in \mathbb{N}.$$

By applying (2.29) and (2.18), we have

$$c \geq \lim_n \Big(\mathbb{E}[\mathbb{E}[X|\pi_n]Y] - \rho(\mathbb{E}[X|\pi_n])\Big) = \mathbb{E}[XY] - \rho(X).$$

Therefore,

$$c \geq \sup_{X \in \mathcal{C}} \{\mathbb{E}[XY] - \rho(X)\}. \tag{2.30}$$

Combining (2.28), and (2.30) is yielding

$$\sup_{X \in \mathcal{L} \cap \mathcal{C}} \{\mathbb{E}[XY] - \rho(X)\} = \sup_{X \in \mathcal{C}} \{\mathbb{E}[XY] - \rho(X)\}.$$

$\square$

2.5 Application

2.5.1 Dilatation Monotonicity and Quasiconvex Law-invariant Functionals

In this section, we present our applications to the study of quasiconvex law-invariant functional (Corollary 27). First let us recall the following result from [63] (Lemma 1.3, Step 2).

Lemma 25. *([63], Lemma 1.3) Let $X \in L^\infty$, $\epsilon > 0$ and $\pi \in \Pi$, there exist $X_1, \dots, X_N \in L^\infty$ that that the same distribution as X and satisfy*

$$\left\| \frac{1}{N}\sum_{i=1}^{N} X_i - \mathbb{E}[X|\pi] \right\|_\infty \leq \epsilon.$$

Using Lemma 17, and Lemma 25, combining with proof of ([17], Proposition 2.2), yields the following construction, which we isolate here.

Lemma 26. *([17], Proposition 2.2) For any $X \in L^1$, there exists $k \in \mathbb{R}_+$ and $X_{n,j}, j = 1, ..., N_n, n \in \mathbb{N}$ such that*

(i) $X_{n,j}\mathbb{1}_{\{|X|\leq n\}}$ *has the same distribution as* $X\mathbb{1}_{\{|X|\leq n\}}$

(ii) $X_{n,j}\mathbb{1}_{\{|X|>n\}} = X\mathbb{1}_{\{|X|>n\}}$ *for each* j, n

(iii) and finally

$$\frac{1}{N_n}\sum_{j=1}^{N_n} X_{n,j} \xrightarrow{a.s.} \mathbb{E}[X|\pi] \quad \textit{and} \quad \left| \frac{1}{N_n}\sum_{j=1}^{N_n} X_{n,j} \right| \leq |X| + |\mathbb{E}[X|\pi]| + k, \quad \textit{for all } n \in \mathbb{N}.$$

Corollary 27. *Let $\mathcal{X}$ be a convex subset of L^1 with the following properties*

(i) $\mathcal{X} + \mathcal{X} \subset \mathcal{X}$,

(ii) $L^\infty \subset \mathcal{X}$ *and* $X\mathbb{1}_A \in \mathcal{X}$ *for each* $X \in \mathcal{X}$ *and* $A \in \mathcal{F}$.

Then for any quasiconvex law-invariant map $\rho : \mathcal{X} \to (-\infty, \infty]$ with the Fatou property, we have that $\overline{\rho} : L^1 \to (-\infty, \infty]$, defined in (2.23), agrees with ρ on $\mathcal{X}$, is quasiconvex, law-invariant, dilatation monotone and $\sigma(L^1, \mathcal{L})$ lower semicontinuous. Moreover, $\overline{\rho}$ preserves monotonicity, convexity, and cash-additivity of ρ.

Proof. First, we will verify that ρ is dilatation monotone, then the result follows by Theorem 22.
Let $X \in \mathcal{X} \subset L^1$ and $\pi \in \Pi$. Pick $X_{n,j}, j = 1, ..., N_n, n \in \mathbb{N}$ as in Lemma 26. Then

$$X_{n,j}\mathbb{1}_{\{|X| \le n\}} \in L^\infty,$$

Since by (i), $\mathcal{X} + \mathcal{X} \subset \mathcal{X}$, thus

$$X_{n,j} = X_{n,j}\mathbb{1}_{\{|X| \le n\}} + X_{n,j}\mathbb{1}_{\{|X| > n\}} = X_{n,j}\mathbb{1}_{\{|X| \le n\}} + X\mathbb{1}_{\{|X| > n\}} \in \mathcal{X}.$$

Also by Lemma 26(ii), we have that $X_{n,j}$ has the same distribution as X, therefore since ρ is law invariant, we get

$$\rho(X) = \rho(X_{n,j}) \quad \text{for each } j, n. \tag{2.31}$$

By the quasi-convexity of ρ and (2.31), we get

$$\rho\left(\frac{1}{N_n}\sum_{j=1}^{N_n} X_{n,j}\right) \le \rho(X).$$

Finally since ρ has the Fatou property, it follows that

$$\rho(\mathbb{E}[X|\pi]) \le \liminf_n \rho\left(\frac{1}{N_n}\sum_{j=1}^{N_n} X_{n,j}\right) \le \rho(X)$$

. □

Note that, Theorem 22 and Corollary 27 improve and extend some recent results in this topic (see, e.g. ([67], Theorem 3.1) and ([52], Theorem 18, Proposition 20)) to domains $\mathcal{X}$ that are not necessarily linear nor rearrangement invariant.

2.5.2 Dilatation Mototonicity of Cash-invariance Hulls

Throughout the financial and insurance literature there is a variety of risk measures or premium principles. However, many of these traditional risk measures fail either to be monotone or cash-invariant. In order to overcome the lack of monotonicity or cash-invariance of risk measures and to provide more useful tools to quantify risks, the notion of monotone and cash-invarianc hulls have been introduced in [27]. This is actually the greatest monotone and cash-invariant function majorized by the (risk) function ρ. In other words, the monotone cash-invariant hull at a given point is nothing else than the optimal objective value of a convex optimization problem. We conclude this Chapter by investigating the dilatation monononicity of cash-invariance hulls.

Definition 28. *For $f : L^1 \to (-\infty, \infty]$, the cash-invariant hull ρ^f of f is defined as follows:*

$$\rho^f(X) := \inf_{s\in\mathbb{R}}\{f(X-s)-s\}, \quad X \in L^1. \tag{2.32}$$

From the definition of ρ^f, it is clear that ρ^f is cash-invariant, and it is easy to see that ρ^f preserves convexity, monotonicity and dilatation monotonicity of f.

In the following result, we show that under a coercive condition, the infimum in (2.32) is attained, and moreover, ρ^f preserves continuity properties of f.

Theorem 29. *Let $f : L^1 \to (-\infty, \infty]$ be dilatation monotone and $\|\cdot\|_1$-lower semicontinuous that satisfies the following coercive condition.*

$$\lim_{|s|\to\infty} f(s) + s = \infty. \tag{2.33}$$

Then $\rho^f : L^1 \to (-\infty, \infty]$ is $\|\cdot\|_1$-lower semicontinuous and

$$\rho^f(X) = \min_{s\in\mathbb{R}}\{f(X-s)-s\}, \quad \textit{for all } X \in L^1.$$

Proof. Pick any $X \in L^1$. If $\rho^f(X) = \infty$, then

$$f(X - s) - s = \infty, \quad \text{for all } s \in \mathbb{R},$$

and thus the infimum in (2.32) is attained at every $s \in \mathbb{R}$. Now assume that $\rho^f(X) < \infty$. Now, we show that, for any $X \in L^1$,

$$\rho^f(X) \neq -\infty. \tag{2.34}$$

Let $(s_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ such that $f(X-s_n)-s_n \to \rho^f(X)$, we show that $(s_n)_{n\in\mathbb{N}}$ is bounded. Suppose that $(s_n)_{n\in\mathbb{N}}$ is unbounded. By passing to a subsequence, we may assume that $|s_n| \to \infty$. Since f is dilatation monotone, we have

$$f(X - s_n) \geq f(\mathbb{E}[X - s_n]). \tag{2.35}$$

Using (2.35), that we get

$$f(X - s_n) - s_n \geq \Big(f(\mathbb{E}[X] - s_n) + (\mathbb{E}[X] - s_n)\Big) - \mathbb{E}[X]. \tag{2.36}$$

Then, we have

$$\rho^f(X) \geq \limsup_n \Big(f(\mathbb{E}[X] - s_n) + (\mathbb{E}[X] - s_n)\Big) - \mathbb{E}[X].$$

Since $X \in L^1$, and $|s_n| \to \infty$, we get

$$|\mathbb{E}[X] - s_n| \geq |s_n| - |\mathbb{E}[X]| \to \infty,$$

Finally (2.33) implies that $\rho^f(X) = \infty$, which is a contradiction.

Thus $(s_n)_{n\in\mathbb{N}}$ is bounded, and by passing to a subsequence, we may assume that $s_n \to s \in \mathbb{R}$. Therefore,

$$X - s_n \xrightarrow{\|\cdot\|_1} X - s.$$

Since f is $\|\cdot\|_1$-lower semicontinuous, we get that

$$f(X-s)-s \leq \liminf_n (f(X-s_n)-s_n) = \rho^f(X).$$

In particular, it follows that $\rho^f(X) = f(X-s)-s$, i.e., the infimum in (2.32) is attained, and consequently, $\rho^f(X) \neq -\infty$.

It remains to be shown that ρ^f is $\|\cdot\|_1$-lower semicontinuous. Fix some $\lambda \in \mathbb{R}$ and let $(X_n)_{n\in\mathbb{N}} \subset L^1, X \in L^1$ be such that

$$X_n \xrightarrow{\|\cdot\|_1} X$$

and $(X_n)_{n\in\mathbb{N}} \subset \{\rho \leq \lambda\}$.

Let $(s_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ be such that $\rho^f(X_n) = f(X_n - s_n) - s_n$. Suppose that $(s_n)_{n\in\mathbb{N}}$ is unbounded. Then we can find a subsequence $(s_{n_k})_{k\in\mathbb{N}}$ such that $|s_{n_k}| \to \infty$. Put

$$t_k := \mathbb{E}[X_{n_k}] - s_{n_k}$$

and note that

$$\mathbb{E}[X_{n_k}] \to \mathbb{E}[X] \in \mathbb{R}, \quad \text{and} \quad |t_k| \geq |s_{n_k}| - |\mathbb{E}[X_{n_k}]| \to \infty.$$

Since f is dilatation monotone, we have that

$$\begin{aligned}
\rho^f(X_{n_k}) &= f(X_{n_k} - s_{n_k}) - s_{n_k} \\
&\geq f(\mathbb{E}[X_{n_k} - s_{n_k}]) - s_{n_k} \\
&\geq f(t_k) + t_k - \mathbb{E}[X_{n_k}].
\end{aligned}$$

By (2.33),

$$f(t_k) + t_k \to \infty,$$

and thus

$$\lim_k \rho^f(X_{n_k}) = \infty,$$

which is a contradiction.

Thus, $(s_n)_{n\in\mathbb{N}}$ is bounded, and we can extract a subsequence $(s_{n_k})_{-k\in\mathbb{N}}$ such that $s_{n_k} \to s \in \mathbb{R}$. Then

$$X_{n_k} - s_{n_k} \xrightarrow{\|\cdot\|_1} X - s$$

Since f is $\|\cdot\|_1$-lower semicontinuous, we have that

$$f(X - s) \leq \liminf_k f(X_{n_k} - s_{n_k})$$

Finally, we get

$$\begin{aligned}
\rho^f(X) &\leq f(X - s) - s \\
&\leq \liminf_k f(X_{n_k} - s_{n_k}) + \lim_k(-s_{n_k}) \\
&= \liminf_k \left(f(X_{n_k} - s_{n_k}) - s_{n_k}\right) \\
&= \liminf_k \rho^f(X_{n_k}) \leq \lambda.
\end{aligned}$$

Therefore, $X \in \{\rho \leq \lambda\}$ and ρ is $\|\cdot\|_1$-lower semicontinuous.

□

The Haezendonck-Goovaerts risk measure is a particular case of the class of cash-invariant hulls ρ^f. Indeed, when X is represent a risk variable in a loss-profit style, by letting $f(X) := \frac{\|X_+\|_p}{(1-\alpha)^{\frac{1}{p}}}$, where $p \geq 1$, then ρ^f is the Haezendonck-Goovaerts risk measure with a power Young function $\varphi(x) = x^p, p \geq 1$ (see, equation (3.6)). Recently, this risk measure has received much attention in actuarial science literature with applications in (re)insurance and portfolio management. In the Next Chapter, we will study various properties of this risk measure.

Chapter 3

The Haezendonck-Goovaerts Risk Measure

In this Chapter, we are interested in studying the Haezendonck-Goovaerts risk measure, which is defined via a power Young function $\Phi(x) = x^p,\ p \geq 1$ with a confident level $\alpha \in (0, 1)$. First, in Section 3.2, we recall some classic results regarding this measure and its optimizer (Orlicz quantile) properties. Our main contribution in this study is Theorem 40 where we prove that for any given $\alpha \in (0, 1)$, the Orlicz quantile for the Haezendonck-Goovaerts risk measure with the power Young function $\Phi(x) = x^2$ is less than $F^{-1}(\alpha)$. Section 3.3.2 is dedicated to investigating the HG risk measure when the risk variable X follows some specific distributions like the uniform, exponential and normal distribution. Moreover, in this Chapter, we provide a methodology to calculate the Haezendonck-Goovaerts risk measure empirically by considering a power Young function $\Phi(x) = x^2$. For practical applications of these results, we refer the reader to the Next Chapter.

3.1 Preliminary Study

3.1.1 Introduction

The Haezendonck-Goovaerts (HG) risk measure, which is defined via a convex Young function Φ and a confidence level parameter α, was first introduced by Haezendonck and Goovaerts in 1982 ([41]) and then re-investigate by Goovaerts et al. in [35, 37]. The HG risk measure can be seen as a generalization of the Average Value at Risk ([61]), and they indeed share some of its well-known properties. As pointed out by Bellini and Rosazza Gianin ([7, 9]), the HG risk measure is a law invariant coherent risk measure.

Recently, the HG risk measure draws attention among researchers as an application in insurance and portfolio management. Numerous studies have been done on this risk measure. For example, a dual representation of this risk measure is given in [7, 9]. Goovaerts et al. in [36] investigate the relationship between the HG risk measure and other risk measures and discuss a generalized version of the HG risk measure. Stability properties of the HG risk measure are studied in [33]. In [8], a computation of the HG measure with a numerical comparison of efficient frontiers between the mean/HG measure and the mean/variance frameworks is investigated. Moreover, this risk measure has been the subject of study in other literature; see, for instance, [3, 36, 53, 65, 66], among others.

Due to the complexity of study the HG measure with the general Young function, the Young function is restricted to be a power Young function $\Phi(x) = x^p,\ p \geq 1$ has been assumed in the literature (see, for instance, [48, 54, 56, 65]). In [65], the asymptotic behaviour of the HG risk measure with a power Young function is investigated. The authors in [54] re-establish the first-order approximate of the HG risk measure in [65]. Under the alternative name of higher-order coherent risk measure, [48, 56] illustrate the advantage of using this risk measure as a risk criterion in portfolio optimization when compared with the Average Value at Risk ($\mathrm{AV@R}_\alpha$). In the following Chapters, we will also focus on the HG risk measure base on a power Young function. More specifically, we are interested in the case of $p = 2$.

3.1.2 The Haezendonck–Goovaerts Risk Measure

We start this section with the HG risk measure definition and then recall some classical results on the HG risk measure. In general, the HG risk measure is naturally defined on Orlicz spaces. This feature makes it more popular in the field of risk measures because of a well-established theory that generalizes the standard L^p theory. Let us recall here some important facts about Young functions and Orlicz Spaces. We are following Edgar and Sucheston ([26]) and Cheridito and Li ([19, 20]) notation in this matter.

Definition 30. *A function $\Phi : [0, +\infty) \to [0, +\infty)$ is called a Young function if it satisfies the following assumptions:*

i) $\lim_{x \to 0^+} \Phi(x) = \Phi(0) = 0$*,* $\lim_{x \to +\infty} \Phi(x) = +\infty$*,*

ii) Φ *is left-continuous, and convex.*

Due to these properties, the Young function Φ is non-decreasing, by which we mean that $\Phi(x) \geq \Phi(y)$ for all $x \geq y$. A Young function Φ is said to be normalized if $\Phi(1) = 1$. Every Young function with $\Phi(1) > 0$ can be normalized by taking $\frac{\Phi(x)}{\Phi(1)}$.

Definition 31. *Let* Φ *be a Young function, then set*

$$L^{\Phi} := \left\{ X : \ \mathbb{E}\left[\Phi\left(\frac{|X|}{\lambda} \right) \right] < +\infty, \quad \textit{for some } \lambda > 0 \right\},$$

is called Orlicz space associated with the Young function Φ*, while the Orlicz heart is its subset and defines as the following*

$$M^{\Phi} := \left\{ X : \ \mathbb{E}\left[\Phi\left(\frac{|X|}{\lambda} \right) \right] < +\infty, \quad \textit{for every } \lambda > 0 \right\}.$$

The Luxemburg norm of $X \in L^{\Phi}$ *is given by*

$$\|X\|_{\Phi} := \inf \left\{ \lambda > 0 : \ \mathbb{E}\left[\Phi\left(\frac{|X|}{\lambda} \right) \right] \leq 1 \right\}.$$

Note that, we always have $L^{\infty} \subset L^{\Phi} \subset L^1$.

Example 32. *Let* $\Phi(x) = x^p, \ p \in [1, \infty)$*, then*

$$M^{\Phi} = L^{\Phi} = L^p.$$

We recall below some basic definitions regarding the HG map by following the style of Bellini and Rosazza Gianin in [7,8]. Let X be a random variable, representing a risk variable in a loss-profit style with distribution function F, then the Orlicz premium principle of risk X is defined as follows.

Definition 33. *Let* Φ *be a normalized Young function. Given* $X \in L^{\Phi}_{+}$ *and* $\alpha \in [0, 1)$*, the Orlicz premium* *principle of risk X is defined as*

$$H_\alpha(X) := \inf\left\{\lambda > 0 : \mathbb{E}\left[\Phi\left(\frac{X}{\lambda}\right)\right] \leq 1-\alpha\right\} = \|X\|_{\Phi_\alpha}, \tag{3.1}$$

where $\Phi_\alpha := \frac{\Phi}{1-\alpha}$.

Definition 34. *For any $X \in L^\Phi$ and $\alpha \in [0,1)$, the Haezendonck-Goovaerts risk is defined as*

$$\pi_{(\Phi,\alpha)}(X) := \inf_{s\in\mathbb{R}} \pi_{(\Phi,\alpha)}(X,s), \tag{3.2}$$

where

$$\pi_{(\Phi,\alpha)}(X,s) := H_\alpha((X-s)_+) + s. \tag{3.3}$$

The actuarial intuition of the HG measure is provided by Bellini and Rosazza Gianin in [8]. If one treats $(X-s)_+$ as the payment for the loss X when applying deductible s, $H_\alpha(.)$ represents the corresponding Orlicz premium of this insurance contract. The minimization construction in the definition minimizes the consumption of the insurer. The minimizer s^* in the optimization problem (3.2) represents the optimal deductible choice from the insurer's point of view. Note that, the parameter α is related to the degree of risk aversion of the insurer.

Remark 35. *For this study, we assume $\alpha \in (0,1)$. Note that, at $\alpha = 0$ the infimum may be attained or unattained (see, [7], Example 15), while for $\alpha \in (0,1)$ it is always attained (see, [9], Proposition 3).*

3.2 The Haezendonck–Goovaerts Risk Measure in L^p

3.2.1 A General Overview

In this Thesis, we restrict the Young function to be a power Young function $\Phi(x) = x^p,\ p \geq 1$. In what follows, we let X be a risk variable distributed by F, and we may assume X belong to the following set,

$$\begin{aligned} \mathcal{Y}_p := \{X : X \text{ be a continuous random variable, } \mathbb{E}[X_+^p] < \infty,\ X \sim F \\ \text{with the endpoints } \tilde{x},\ \tilde{y} \text{ such that } -\infty \leq \tilde{y} < \tilde{x} \leq +\infty\} \end{aligned} \quad (3.4)$$

We denote $\pi_{(p,\alpha)}(X)$, the HG risk measure of X with a power Young function $\Phi(x) = x^p,\ p \geq 1$, at a given level of $\alpha \in (0,1)$.

We recall that when $p = 1$, the HG risk measure is corresponding to the Young function $\Phi(x) = x$ is the $\mathrm{AV@R}_\alpha$. If $\Phi(x) = x^p,\ p \geq 1$, then for any $X \in L^\Phi = L^p$, using Definition 33, we get

$$H_\alpha(X) = \left(\frac{\mathbb{E}[X^p]}{1-\alpha}\right)^{\frac{1}{p}}, \quad \alpha \in (0,1). \quad (3.5)$$

Then the HG risk measure is corresponding to the Young function $\Phi(x) = x^p,\ p \geq 1$ is defined by putting more weight on the right tail of X as the following

$$\pi_{(p,\alpha)}(X) := \inf_{s \in \mathbb{R}} \left\{ \left(\frac{\mathbb{E}[(X-s)_+^p]}{1-\alpha}\right)^{\frac{1}{p}} + s \right\}, \quad p \geq 1. \quad (3.6)$$

It is usually hard to obtain an explicit expression for the HG risk measure with the general Young function. Here below, we provide an important result which we use in our study. In the case p=1, Rockafellar and Uryasev ([62], Theorem 10) show that in this situation, the infimum in the equation (3.6) is always attained. More precisely, the set of minimizers coincides with the usual α-quantiles and have a representation though the optimization problem (3.6) when $p = 1$. For a power Young function $\Phi(x) = x^p$ with $p > 1$, Tang and Yang ([65], Theorem 2.1), and for more general case, $X \in M^\Phi$, where Φ is a strictly convex function, Bellini and Rosazza Gianin ([9], Proposition 6) are established a general result for the HG risk measure. In the following, we summarized those results when $\Phi(x) = x^p,\ p \geq 1$.

Theorem 36. *Let the Young function be $\Phi(x) = x^p$ for some $p \geq 1$ and let $X \in \mathcal{Y}_p$ be a risk variable. Then*

(i) If $p = 1$, the Average Value at Risk for X is equal to

$$\begin{aligned} \mathrm{AV@R}_\alpha(X) :&= \inf_{s\in\mathbb{R}} \left\{ \frac{1}{1-\alpha}\mathbb{E}[(X-s)_+] + s \right\} \\ &= \frac{1}{1-\alpha}\mathbb{E}\left[\left(X - F^{-1}(\alpha)\right)_+ \right] + F^{-1}(\alpha), \quad \alpha \in (0,1). \end{aligned} \tag{3.7}$$

(ii) For any $p > 1$, the Haezendonck–Goovaerts risk measure for X is equal to

$$\begin{aligned} \pi_{(p,\alpha)}(X) :&= \inf_{s\in\mathbb{R}} \left\{ \left(\frac{\mathbb{E}[(X-s)_+^p]}{1-\alpha} \right)^{\frac{1}{p}} + s \right\} \\ &= \left(\frac{\mathbb{E}[(X-s^*)_+^p]}{1-\alpha} \right)^{\frac{1}{p}} + s^*, \quad \alpha \in (0,1), \end{aligned} \tag{3.8}$$

where $s^* := s^*_{p,\alpha} \in (-\infty, \tilde{x})$ is the unique solution to the equation

$$\xi(s,\alpha) := \left(\mathbb{E}[(X-s)_+^{p-1}] \right)^{\frac{1}{p-1}} - (1-\alpha)^{\frac{1}{p(p-1)}} \left(\mathbb{E}[(X-s)_+^p] \right)^{\frac{1}{p}} = 0. \tag{3.9}$$

The minimizer $s^*_{p,\alpha}$ sometimes is called L^p-quantile or Orlicz quantile. Obviously, the value of s^* in (3.8) depends on the choice of a Young function $x^p,\ p > 1$ and the parameter α and the random variable X. In order to avoid heavy notations, we will write s^* instead of $s^*_{p,\alpha}(X)$ whenever there is no confusion.

We should emphasize that; actually, the formulas (3.8) holds for both a discrete and a continuous random variable X (we refer reader to [15] for the general case.). Using this fact, and with the aid of Theorem 36, we shall establish an empirical method to estimate the value of the HG risk measure when p=2 in Section 3.4.

3.2.2 A Connection Between $\pi_{(2,\alpha)}(.)$ and the $\mathrm{AV@R}(.)$ and More

In [48] is shown that $\pi_{(p,\alpha)}(.)$ behaves in a similar way in properties to $\mathrm{AV@R}(.)$. In the following, we restate and highlight some results from [9] and [48], base on a power Young function $x^p, p \geq 1$.

Theorem 37. *For any $p \in [1, \infty)$, and given $\alpha \in (0,1)$, the HG map $\pi_{(p,\alpha)} : L^p \to \mathbb{R}$, defined as in (3.6) satisfies the following properties:*

(i) $\pi_{(p,\alpha)}$ is coherent risk measure[1].

(ii) $\pi_{(p,\alpha)}$ is law-invariant.

(iii) $\pi_{(p,\alpha)}$ is dilatation monotone.

(iv) With fixed X and α, then $\pi_{(p,\alpha)}$ is non-decreasing function with respect to p.

(v) $\pi_{(p,\alpha)}$ is satisfies the Fatou property.

(vi) Given p and $X \in L^p$, then the map $\pi_{(p,\alpha)}$ is non-decreasing function with respect to α.

Proof. (i) It is easy to check that the map $\pi_{(p,\alpha)}$ is a coherent risk measure.

Cash-invariant: For any $X \in L^p$ and $l \in \mathbb{R}$, we have

$$\begin{aligned}
\pi_{(p,\alpha)}(X+l) &= \inf_{s\in\mathbb{R}} \left\{ \left(\frac{\mathbb{E}[(X+l-s)_+^p]}{1-\alpha} \right)^{\frac{1}{p}} + s \right\} \\
&= \inf_{s-l\in\mathbb{R}} \left\{ \left(\frac{\mathbb{E}[(X-(s-l))_+^p]}{1-\alpha} \right)^{\frac{1}{p}} + (s-l) \right\} + l \\
&= \pi_{(p,\alpha)}(X) + l.
\end{aligned}$$

[1] Differently from most of financial literature, as it is consequence of the profit/loss assumption.

Monotonicity: Let $X, Y \in L^p$ such that $X \leq Y$. Since $\mathbb{E}[(X-s)_+^p] \leq \mathbb{E}[(Y-s)_+^p]$, we get

$$\begin{aligned}
\pi_{(p,\alpha)}(X) &= \inf_{s\in\mathbb{R}} \left\{ \left(\frac{\mathbb{E}[(X-s)_+^p]}{1-\alpha} \right)^{\frac{1}{p}} + s \right\} \\
&\leq \inf_{s\in\mathbb{R}} \left\{ \left(\frac{\mathbb{E}[(Y-s)_+^p]}{1-\alpha} \right)^{\frac{1}{p}} + s \right\} \\
&= \pi_{(p,\alpha)}(Y).
\end{aligned}$$

Positive homogeneity: For any $\lambda > 0$ and $X \in L^p$, we have

$$\begin{aligned}
\pi_{(p,\alpha)}(\lambda X) &= \inf_{s\in\mathbb{R}} \left\{ \left(\frac{\mathbb{E}[(\lambda X-s)_+^p]}{1-\alpha} \right)^{\frac{1}{p}} + s \right\} \\
&= \inf_{s\in\mathbb{R}} \left\{ \lambda \left(\frac{\mathbb{E}[(X-\frac{s}{\lambda})_+^p]}{1-\alpha} \right)^{\frac{1}{p}} + \lambda \frac{s}{\lambda} \right\} \\
&= \lambda \inf_{\frac{s}{\lambda}\in\mathbb{R}} \left\{ \left(\frac{\mathbb{E}[(X-\frac{s}{\lambda})_+^p]}{1-\alpha} \right)^{\frac{1}{p}} + \frac{s}{\lambda} \right\} \\
&= \lambda \pi_{(p,\alpha)}(X).
\end{aligned}$$

when $\lambda = 0$, for any $X \in L^p$, Since

$$\begin{aligned}
\pi_{(p,\alpha)}(0) &= \inf_{s\in\mathbb{R}} \left\{ \left(\frac{\mathbb{E}[(-s)_+^p]}{1-\alpha} \right)^{\frac{1}{p}} + s \right\} \\
&= \inf \left\{ \inf_{s<0} \left\{ \frac{-s}{(1-\alpha)^{\frac{1}{p}}} + s \right\}, \inf_{s\geq 0} s \right\} = 0.
\end{aligned}$$

For any $\lambda \geq 0$, we get,

$$\pi_{(p,\alpha)}(\lambda X) = \lambda \pi_{(p,\alpha)}(X).$$

Subadditivity: For any $X, Y \in L^p$, we have

$$\begin{aligned}
\pi_{(p,\alpha)}(X+Y) &= \inf_{s\in\mathbb{R}}\left\{\left(\frac{\mathbb{E}[(X+Y-s)_+^p]}{1-\alpha}\right)^{\frac{1}{p}} + s\right\} \\
&= \inf_{s_1,s_2\in\mathbb{R}}\left\{\left(\frac{\mathbb{E}[(X-s_1+Y-s_2)_+^p]}{1-\alpha}\right)^{\frac{1}{p}} + s_1 + s_2\right\} \\
&\leq \inf_{s_1,s_2\in\mathbb{R}}\left\{\left(\frac{\mathbb{E}[(X-s_1)_+^p]}{1-\alpha}\right)^{\frac{1}{p}} + s_1 + \left(\frac{\mathbb{E}[(Y-s_2)_+^p]}{1-\alpha}\right)^{\frac{1}{p}} + s_2\right\} \\
&\leq \inf_{s_1\in\mathbb{R}}\left\{\left(\frac{\mathbb{E}[(X-s_1)_+^p]}{1-\alpha}\right)^{\frac{1}{p}} + s_1\right\} + \inf_{s_2\in\mathbb{R}}\left\{\left(\frac{\mathbb{E}[(Y-s_2)_+^p]}{1-\alpha}\right)^{\frac{1}{p}} + s_2\right\} \\
&= \pi_{(p,\alpha)}(X) + \pi_{(p,\alpha)}(Y).
\end{aligned}$$

Thus, $\pi_{(p,\alpha)}$ is a coherent risk measure.

(ii) *Law-invariant*: Let $X, Y \in L^p$ such that $X \sim Y$. Then

$$\mathbb{E}[(X-t)_+^p] = \mathbb{E}[(Y-t)_+^p], \quad \text{for all } t \in \mathbb{R}.$$

And

$$\begin{aligned}
\pi_{(p,\alpha)}(X) &= \inf_{s\in R}\left\{\left(\frac{\mathbb{E}[(X-s)_+^p]}{1-\alpha}\right)^{\frac{1}{p}} + s\right\} \\
&\leq \left(\frac{\mathbb{E}[(X-t)_+^p]}{1-\alpha}\right)^{\frac{1}{p}} + t \\
&= \left(\frac{\mathbb{E}[(Y-t)_+^p]}{1-\alpha}\right)^{\frac{1}{p}} + t \\
&\leq \pi_{(p,\alpha)}(Y).
\end{aligned}$$

Similarly, $\pi_{(p,\alpha)}(Y) \leq \pi_{(p,\alpha)}(X)$. Therefore, if $X \sim Y$ then

$$\pi_{(p,\alpha)}(Y) = \pi_{(p,\alpha)}(X).$$

(iii) *Dilatation monotonicity*: For any $X \in L^p$, it is easy to see

$$\pi_{(p,\alpha)}(\mathbb{E}[X|\pi]) \leq \pi_{(p,\alpha)}(X).$$

Indeed, by the conditional Jensen's inequality and the contraction properties of $\|.\|_p$-norm , we get

$$\begin{aligned}
\pi_{(p,\alpha)}(\mathbb{E}[X|\pi]) &= \inf_{s\in\mathbb{R}} \left\{ \left(\frac{\mathbb{E}[(\mathbb{E}[X|\pi] - s)_+^p]}{1-\alpha} \right)^{\frac{1}{p}} + s \right\} \\
&= \inf_{s\in\mathbb{R}} \left\{ \left(\frac{\mathbb{E}[(\mathbb{E}[X - s|\pi])_+^p]}{1-\alpha} \right)^{\frac{1}{p}} + s \right\} \\
&\leq \inf_{s\in\mathbb{R}} \left\{ \left(\frac{\mathbb{E}[(X - s)_+^p]}{1-\alpha} \right)^{\frac{1}{p}} + s \right\} \\
&= \pi_{(p,\alpha)}(X).
\end{aligned}$$

(iv) Let assume $p \geq 1$ and $p < r$, then

$$\left(\mathbb{E}[(X-s)_+^p]\right)^{\frac{1}{p}} \leq \left(\mathbb{E}[(X-s)_+^r]\right)^{\frac{1}{r}}, \quad \text{for any } p < r \text{ and } X \in L^r.$$

Then clearly, for any $\alpha \in (0,1)$, $X \in L^r$

$$\begin{aligned}
\pi_{(p,\alpha)}(X) &= \inf_{s\in\mathbb{R}} \left\{ \left(\frac{\mathbb{E}[(X - s)_+^p]}{1-\alpha} \right)^{\frac{1}{p}} + s \right\} \\
&\leq \inf_{s\in\mathbb{R}} \left\{ \left(\frac{\mathbb{E}[(X - s)_+^r]}{1-\alpha} \right)^{\frac{1}{r}} + s \right\} \\
&= \pi_{(r,\alpha)}(X).
\end{aligned}$$

Thus $\pi_{(p,\alpha)}(.)$ is a non-decreasing map respect to p.

(v) This is easy to see from Theorem 29.

(vi) Let assume $\alpha_1, \alpha_2 \in (0,1)$ and $\alpha_1 < \alpha_2$. Then

$$\frac{\mathbb{E}[(X-s)_+^p]^{\frac{1}{p}}}{1-\alpha_1} < \frac{\mathbb{E}[(X-s)_+^p]^{\frac{1}{p}}}{1-\alpha_2}. \tag{3.10}$$

Thus, for any $\alpha_1 < \alpha_2$,

$$\pi_{(p,\alpha_1)}(X) \leq \pi_{(p,\alpha_2)}(X), \quad \text{for any } X \in L^p.$$

□

By this Theorem, it is easy to see that the $\mathrm{AV@R}_\alpha(.)$ (case p=1) is the smallest HG risk measure among all families of coherent risk measure $\pi_{(p,\alpha)}(.)$.

3.3 An Approximation Method

As we indicated before, in this thesis, our main focus is at the case when $p = 2$. For simplification of our study when the Young function is $\varphi(x) = x^2$, we restatement Theorem 36 (ii). By plugging $p = 2$ into the equation (3.9), then the unique solution $s^*_{p,\alpha}$ satisfies in the following equation,

$$\left(\mathbb{E}[(X-s)_+]\right)^2 - (1-\alpha)\mathbb{E}\left[(X-s)_+^2\right] = 0. \tag{3.11}$$

Using (3.11), for any $\alpha \in (0,1)$ and $X \in \mathcal{Y}_2$, we get

$$\begin{aligned}\pi_{(2,\alpha)}(X) &= \left(\frac{\mathbb{E}[(X-s^*)_+^2]}{1-\alpha}\right)^{\frac{1}{2}} + s^* \\ &= \frac{\mathbb{E}[(X-s^*)_+]}{1-\alpha} + s^*.\end{aligned} \tag{3.12}$$

Implementation-wise, when p=2, the $\pi_{(2,\alpha)}(.)$ can be incorporated into a mathematical programming problem via the second-order cone constraints. The second-order cone programming is a well-developed topic in the field of convex optimization (see, for example, [4]). Several commercial off-the-shelf software packages are available for solving this kind of optimization, for an instant, *MOSEK* and *NAG Numerical Library*.

3.3.1 A General Approximation

In the following theorem, we provide a list of properties that the minimizer $s^*_{p,\alpha}(.)$ holds. Note that under the name of Orlicz quantile, some of these properties such as (i)-(iii) are investigated when Φ is the general Young function (see, [9]).

Theorem 38. *For any $p \in [1, \infty)$, and any given $\alpha \in (0, 1)$. Let $X \in \mathcal{Y}_p$ and $s^*_{p,\alpha}(.)$ be the unique solution satisfies in the equation (3.9), then*

(i) $s^*_{p,\alpha}(c) = c$, so the minimizer s^* is normalized, $s^*_{p,\alpha}(0) = 0$.

(ii) For a fixed $\alpha \in (0, 1)$ and $p \geq 1$, the minimizer $s^*_{p,\alpha}(.)$ is cash-invariant, and positively homogeneous.

(iii) The minimizer $s^*_{p,\alpha}(.)$ is law invariance.

(iv) Let $p \geq 1$ be fixed, then $s^*_{p,\alpha}(.)$ is a non-decreasing function respect to α.

Proof. (i) By Theorem 37, $\pi_{(p,\alpha)}$ is cash-invariant and $\pi_{(p,\alpha)}(0) = 0$. Now, let $X = c$ and $s^* := s^*_{p,\alpha}(c)$, then we obtain,

$$\pi_{(p,\alpha)}(c) = c = \frac{(c - s^*)_+}{(1-\alpha)^{\frac{1}{p}}} + s^*. \tag{3.13}$$

Thus, $s^* = c$, and obviously if $c = 0$, then $s^*_{p,\alpha}(0) = 0$.

(ii) *Cash invariant:* Let $l \in \mathbb{R}$, $s^*_{X+l} := s^*_{p,\alpha}(X + l)$ then

$$\begin{aligned}\pi_{(p,\alpha)}(X + l) &= \left(\frac{\mathbb{E}[(X - (s^*_{X+l} - l))^p_+]}{1-\alpha}\right)^{\frac{1}{p}} + s^*_{X+l} - l + l \\ &= \left(\frac{\mathbb{E}[(X - t^*)^p_+]}{1-\alpha}\right)^{\frac{1}{p}} + t^* + l,\end{aligned}$$

where $t^* := s^*_{X+l} - l$.

By cash-invariant of $\pi_{(p,\alpha)}$, we get

$$\left(\frac{\mathbb{E}[(X-t^*)_+^p]}{1-\alpha}\right)^{\frac{1}{p}} + t^* + l = \left(\frac{\mathbb{E}[(X-s_X^*)_+^p]}{1-\alpha}\right)^{\frac{1}{p}} + s_X^* + l. \tag{3.14}$$

And finally since s^* is unique, we get $s_X^* = t^* = s_{X+l}^* - l$.

Positively homogeneous: For any $X \in L^p$, and $\lambda > 0$, and by homogeneity of $\pi_{(p,\alpha)}$, we get

$$\begin{aligned}
\lambda\left(\left(\frac{\mathbb{E}[(X-s^*)_+^p]}{1-\alpha}\right)^{\frac{1}{p}} + s^*\right) &= \left(\frac{\mathbb{E}[(\lambda X-t^*)_+^p]}{1-\alpha}\right)^{\frac{1}{p}} + t^* \\
&= \lambda\left(\frac{\mathbb{E}[(X-\frac{t^*}{\lambda})_+^p]}{1-\alpha}\right)^{\frac{1}{p}} + \lambda\frac{t^*}{\lambda},
\end{aligned}$$

where t^*, s^* are minimizers for $\pi_{(p,\alpha)}(\lambda X)$, and $\pi_{(p,\alpha)}(X)$ respectively.

Since $\lambda \neq 0$, we get

$$\left(\frac{\mathbb{E}[(X-s^*)_+^p]}{1-\alpha}\right)^{\frac{1}{p}} + s^* = \left(\frac{\mathbb{E}[(X-\frac{t^*}{\lambda})_+^p]}{1-\alpha}\right)^{\frac{1}{p}} + \frac{t^*}{\lambda}.$$

Finally since s^* is unique, therefore $\lambda s^* = t^*$.

If $\lambda = 0$, by (i), clearly $\lambda s^* = t^*$. Thus

$$\lambda s_{p,\alpha}(X) = s_{p,\alpha}(\lambda X), \quad \text{for any } \lambda \geq 0.$$

(iii) Let $X, Y \in L^p$ and $X \sim Y$, then since $\pi_{(p,\alpha)}$ is law invariant and

$$\mathbb{E}[(X-s^*)_+^p] = \mathbb{E}[(Y-s^*)_+^p].$$

Thus, we obtain

$$\left(\frac{\mathbb{E}[(X-s^*)_+^p]}{1-\alpha}\right)^{\frac{1}{p}} + s^* = \left(\frac{\mathbb{E}[(Y-s^*)_+^p]}{1-\alpha}\right)^{\frac{1}{p}} + s^*$$
$$= \left(\frac{\mathbb{E}[(Y-t^*)_+^p]}{1-\alpha}\right)^{\frac{1}{p}} + t^*.$$

and it follows immediately $s^* = t^*$, as s^* is unique.

(iv) Let $p \geq 1$ be fixed and lets assume $s_1^* := s_{p,\alpha_1}^*(X)$ and $s_2^* := s_{p,\alpha_2}^*(X)$ be solutions to the equation (3.9) at α_1 and α_2 respectively where $\alpha_1 < \alpha_2$. Using the equation (3.8) for the pair (s_2^*, α_2), we get

$$\pi_{(p,\alpha_2)}(X) = \left(\frac{\mathbb{E}[(X-s_2^*)_+^p]}{(1-\alpha_2)}\right)^{\frac{1}{p}} + s_2^* \leq \left(\frac{\mathbb{E}[(X-s_1^*)_+^p]}{(1-\alpha_2)}\right)^{\frac{1}{p}} + s_1^*. \tag{3.15}$$

By rearranging (3.15), we have

$$\left(\frac{\mathbb{E}[(X-s_2^*)_+^p]}{(1-\alpha_2)}\right)^{\frac{1}{p}} - \left(\frac{\mathbb{E}[(X-s_1^*)_+^p]}{(1-\alpha_2)}\right)^{\frac{1}{p}} \leq s_1^* - s_2^*. \tag{3.16}$$

Similarly, using (3.8) for the pair (s_1^*, α_1), we get

$$\left(\frac{\mathbb{E}[(X-s_2^*)_+^p]}{(1-\alpha_1)}\right)^{\frac{1}{p}} - \left(\frac{\mathbb{E}[(X-s_1^*)_+^p]}{(1-\alpha_1)}\right)^{\frac{1}{p}} \geq s_1^* - s_2^*. \tag{3.17}$$

By combining (3.16) and (3.17), we obtain

$$(s_1^* - s_2^*)(1-\alpha_1)^{\frac{1}{p}} \leq \left(\mathbb{E}[(X-s_2^*)_+^p]\right)^{\frac{1}{p}} - \left(\mathbb{E}[(X-s_1^*)_+^p]\right)^{\frac{1}{p}} \leq (s_1^* - s_2^*)(1-\alpha_2)^{\frac{1}{p}}.$$

Now, let assume $s_{p,\alpha}^*$ be a non-increasing function of α. Therefore for any $\alpha_1 < \alpha_2$, $s_1^* > s_2^*$. Then, since $s_1^* - s_2^* > 0$, we get

$$(1-\alpha_1)^{\frac{1}{p}} \leq (1-\alpha_2)^{\frac{1}{p}},$$

which is contradiction with the assumption of $\alpha_1 < \alpha_2$. □

Remark 39.

(i) *Note that, unfortunately, $s^*_{p,\alpha}(X)$ fails to be monotone (increasing) in X, i.e. $X \leq Y$ a.s. does not necessarily implies $s^*_{p,\alpha}(X) \leq s^*_{p,\alpha}(Y)$ (see, Example 45).*

(ii) *Also, $s^*_{p,\alpha}(X)$ can be negative even if X is a non-negative random variable. It seems that this fact is neglected in [65], Theorem 7.1, when $\Phi(x) = x^p, p \geq 1$ (see, Proposition 46).*

The following theorem is our main result in this Chapter (for the first part of result, we also refer the reader to [40]). In this theorem, we show that for any given $\alpha \in (0,1)$, $F^{-1}(\alpha)$ is an upper bound for the minimizer $s^*_{2,\alpha}$.

Theorem 40. *Let $\Phi(x) = x^2$ and $X \in \mathcal{Y}_2$. Then at a given confidence level $\alpha \in (0,1)$, the minimizer $s^* := s^*_{2,\alpha} \in (-\infty, F^{-1}(1))$ satisfies in (3.11) is less than $F^{-1}(\alpha)$.*

Moreover, for any $\alpha \leq \alpha_0$ with

$$\frac{1}{\alpha_0} = \frac{(\mathbb{E}[X] - \tilde{y})^2}{\sigma^2(X)} + 1,$$

$s^* \leq \tilde{y}$, where $\tilde{y}$ is the lower endpoint of X and we have

$$\pi_{(2,\alpha)}(X) = \mathbb{E}[X] + \sigma(X)\sqrt{\frac{\alpha}{1-\alpha}}, \quad s^* = \mathbb{E}[X] - \sigma(X)\sqrt{\frac{1-\alpha}{\alpha}}. \tag{3.18}$$

Proof. Let $p \geq 1$ be an integer, then

$$\begin{aligned}
\mathbb{E}[(X-s)_+^p] &= \int_s^{\tilde{x}} (X-s)^p dF(X) \\
&= \int_s^{\tilde{x}} \sum_{k=0}^{p} \binom{p}{k} (-1)^{p-k} s^{p-k} X^k dF(X) \\
&= \sum_{k=0}^{p} \binom{p}{k} (-1)^{p-k} s^{p-k} \int_s^{\tilde{x}} X^k dF(X).
\end{aligned} \tag{3.19}$$

For simplicity, we denote

$$\bar{\mu}_{k,s}(F) := \int_s^{\tilde{x}} X^k dF(X).$$

Applying (3.19) for $p = 1, 2$, we get the following equations

$$\mathbb{E}[(X - s)_+] = \bar{\mu}_{1,s}(F) - s\bar{F}(s), \tag{3.20}$$

$$\mathbb{E}[(X - s)_+^2] = \bar{\mu}_{2,s}(F) - 2s\bar{\mu}_{1,s}(F) + s^2\bar{F}(s). \tag{3.21}$$

where $\bar{F}$ is the survival function.

By substituting (3.20), and (3.21) in the equation (3.11), we obtain

$$\begin{aligned}
1 - \alpha &= \frac{\left(\bar{\mu}_{1,s}(F) - s\bar{F}(s)\right)^2}{\bar{\mu}_{2,s}(F) - 2s\bar{\mu}_{1,s}(F) + s^2\bar{F}(s)} \\
&= \frac{\left(\bar{\mu}_{1,s}(F) - s\bar{F}(s)\right)^2}{\bar{F}(s)\left(s^2 - 2s\frac{\bar{\mu}_{1,s}(F)}{\bar{F}(s)} + \frac{\bar{\mu}_{2,s}(F)}{\bar{F}(s)}\right)} \\
&= \frac{\bar{F}(s)^2\left(\frac{\bar{\mu}_{1,s}(F)}{\bar{F}(s)} - s\right)^2}{\bar{F}(s)\left(\left(\frac{\bar{\mu}_{1,s}(F)}{\bar{F}(s)} - s\right)^2 + \frac{\bar{\mu}_{2,s}(F)}{\bar{F}(s)} - \left(\frac{\bar{\mu}_{1,s}(F)}{\bar{F}(s)}\right)^2\right)}.
\end{aligned}$$

A rearrangement of the terms leads us to

$$\left(\bar{F}(s) - (1 - \alpha)\right)\left(\frac{\bar{\mu}_{1,s}(F)}{\bar{F}(s)} - s\right)^2 = \frac{(1 - \alpha)}{\bar{F}(s)}\sigma_s^2\left(\frac{\bar{\mu}_{1,s}(F)}{\bar{F}(s)}\right), \tag{3.22}$$

where

$$\sigma_s^2\left(\frac{\bar{\mu}_{1,s}(F)}{\bar{F}(s)}\right) := \mathbb{E}\left[\left(X - \frac{\bar{\mu}_{1,s}(F)}{\bar{F}(s)}\right)^2 \mathbb{1}_{X \geq s}\right] = \bar{\mu}_{2,s}(F) - \frac{\left(\bar{\mu}_{1,s}(F)\right)^2}{\bar{F}(s)}.$$

By Theorem 36, for a fix α, s^* is the unique solution to the equation (3.22). This equation has solution if

$$\left(\bar{F}(s^*) - (1 - \alpha)\right) > 0.$$

Therefore, $s^* < F^{-1}(\alpha)$.

In addition, let assume $s^* \leq \tilde{y}$, where $\tilde{y}$ is the lower endpoint of X

$$\begin{aligned}\mathbb{E}[(X-s)_+^p] &= \int_{\tilde{y}}^{\tilde{x}} (X-s)^p dF(X) \\ &= \int_{\tilde{y}}^{\tilde{x}} \sum_{k=0}^{p} \binom{p}{k} (-1)^{p-k} s^{p-k} X^k dF(X) \\ &= \sum_{k=0}^{p} \binom{p}{k} (-1)^{p-k} s^{p-k} \mathbb{E}[X^k]. \end{aligned} \tag{3.23}$$

With similar processes, we have

$$(\mathbb{E}[X] - s^*)^2 = (1-\alpha)\left((\mathbb{E}[X] - s^*)^2 + \sigma^2(X)\right). \tag{3.24}$$

Note that if $s^* = \tilde{y}$ at the level of α_0, solving (3.24) for α_0, we get

$$\alpha_0 = \frac{\sigma^2(X)}{(\mathbb{E}[X] - s^*)^2 + \sigma^2(X)}. \tag{3.25}$$

By Theorem 38 (iv), for any $\alpha < \alpha_0$, we have $s^* := s^*_\alpha \leq s^*_{\alpha_0} := \tilde{y}$. Thus, by (3.24), we get

$$s^* = \mathbb{E}[X] - \sigma(X)\sqrt{\frac{1-\alpha}{\alpha}}, \quad \text{for any } \alpha \leq \alpha_0.$$

Finally, using (3.12), we obtain

$$\begin{aligned}\pi_{(2,\alpha)}(X) &= \left(\frac{\mathbb{E}[(X-s^*)_+^2]}{1-\alpha}\right)^{\frac{1}{2}} + s^* \\ &= \frac{\mathbb{E}[(X-s^*)_+]}{1-\alpha} + s^* \\ &= \frac{\mathbb{E}[X] - s^*}{1-\alpha} + s^* \\ &= \mathbb{E}[X] + \sigma(X)\sqrt{\frac{\alpha}{1-\alpha}}.\end{aligned}$$

□

Proposition 41. *Let $X \in \mathcal{Y}_2$ be risk variable with $\pi_{(2,\alpha)}(X)$. Then there exist α_0, such that $\alpha_0 = \frac{\sigma^2(X_+)}{\mathbb{E}[X_+^2]}$ and* *for any* $\alpha > \alpha_0$, $s_\alpha^* := s_{2,\alpha}^* \in \big[0, F^{-1}(\alpha)\big)$. *Moreover,*

$$0 \leq \pi_{(2,\alpha)}(X) - AV@R_\alpha(X) < \frac{F^{-1}(\alpha)}{1-\alpha}.$$

Proof. Let $X \in \mathcal{Y}_2$. By letting $s = 0$ in the equation (3.11), at the α_0 level , we get

$$\frac{\left(\mathbb{E}[X_+]\right)^2}{\mathbb{E}[X_+^2]} = 1 - \alpha_0.$$

Solving the above equation for α_0, we have

$$\alpha_0 = \frac{\mathbb{E}[X_+^2] - \left(\mathbb{E}[X_+]\right)^2}{\mathbb{E}[X_+^2]} = \frac{\sigma^2(X_+)}{\mathbb{E}[X_+^2]}.$$

Therefore, $s^* = 0$ is the optimizer for the optimization problem (3.11) at $\alpha_0 = \frac{\sigma^2(X_+)}{\mathbb{E}[X_+^2]}$. Using Theorem 38 (iv), s_α^* is non-decreasing function respect to α. Therefore,

$$s_\alpha^* \geq s_{\alpha_0}^* = 0, \quad \text{for any } \alpha > \alpha_0.$$

Finally, by combining the above result and Theorem 40, we obtain

$$0 \leq s_\alpha^* < F^{-1}(\alpha), \quad \text{for any } \alpha > \alpha_0.$$

By Theorem 37 (iv), for any fixed α, $\mathrm{AV@R}_\alpha(X) \leq \pi_{(2,\alpha)}(X)$. Then by (3.12) and (3.7) for any $\alpha > \alpha_0$, we get

$$\begin{aligned}
\mathrm{AV@R}_\alpha(X) \leq \pi_{(2,\alpha)}(X) &\leq s^*_\alpha + \frac{\mathbb{E}[(X - s^*_\alpha)_+]}{1-\alpha} \\
&< F^{-1}(\alpha) + \frac{\mathbb{E}[(X - s^*_\alpha)_+]}{1-\alpha} \\
&\leq \mathrm{AV@R}_\alpha(X) + \frac{\mathbb{E}[(F^{-1}(\alpha) - s^*_\alpha)_+]}{1-\alpha} \\
&\leq \mathrm{AV@R}_\alpha(X) + \frac{\mathbb{E}[(F^{-1}(\alpha) - s^*_\alpha)]}{1-\alpha} \\
&= \mathrm{AV@R}_\alpha(X) + \frac{F^{-1}(\alpha)}{1-\alpha} - \frac{s^*_\alpha}{1-\alpha}.
\end{aligned}$$

Since $0 \leq s^*_\alpha < F^{-1}(\alpha)$, we get the desire result.

$$0 \leq \pi_{(2,\alpha)}(X) - \mathrm{AV@R}_\alpha(X) < \frac{F^{-1}(\alpha)}{1-\alpha}, \quad \text{for any } \alpha > \alpha_0.$$

$\square$

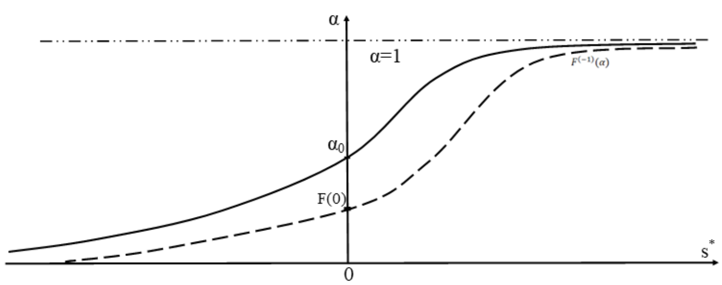


Figure 3.1: This figure illustrates s^* value (x-axis) while α is changing. The solid curve is the optimizer s^*, when p=2 and the dash line is $F^{-1}(\alpha)$ as a minimizer of $\mathrm{AV@R}_\alpha$.

3.3.2 Some Examples

The main ingredient in achieving the objective function (3.12) rests on finding the solution to the equation (3.11). In this respect, we concentrate on the case in which the risk variable X follows the uniform, exponential, and normal distributions, to find the solution $s^*_{p,\alpha}(X)$ for the optimization problem (3.12). Note that we verify the following results independently of Theorem 40.

3.3.2.1 Uniform Distribution

We start with a motivated and easy example by considering the uniform distribution and we obtain analytical expression for the $\pi_{(2,\alpha)}$ (Proposition 42) . Assume $X \sim \mathcal{U}(a,b)$, recall that uniform distribution parameters have range $-\infty < a < b < \infty$ where pdf and cdf are given by,

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \le x \le b, \\ 0, & o.w. \end{cases}, \quad F(x) = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a} & a \le x \le b, \\ 1, & x > b. \end{cases}$$

Its inverse is

$$F^{-1}(\alpha) = a + \alpha(b-a), \quad \text{for } \alpha \in (0,1). \tag{3.26}$$

Proposition 42. *If $\Phi(x) = x^2$, and $X \sim \mathcal{U}(a,b)$, then*

$$\pi_{(2,\alpha)}(X) = \begin{cases} \frac{1}{2}(b+a) + \frac{1}{2}(b-a)\sqrt{\frac{\alpha}{3(1-\alpha)}}, & 0 < \alpha \le \frac{1}{4}, \\ b - \frac{4}{9}(b-a)(1-\alpha), & \frac{1}{4} \le \alpha < 1. \end{cases} \tag{3.27}$$

Proof. Let assume $s \le a$, then

$$\begin{aligned} \mathbb{E}[(X-s)_+^p] &= \int_a^b (X-s)^p f(X)dX = \frac{1}{b-a}\int_a^b (X-s)^p dX \\ &= \frac{(b-s)^{p+1} - (a-s)^{p+1}}{(p+1)(b-a)}. \end{aligned}$$

By substituting in (3.11), when p=2, we have

$$1-\alpha = \frac{\frac{\left((b-s)^2-(a-s)^2\right)^2}{4(b-a)^2}}{\frac{(b-s)^3-(a-s)^3}{3(b-a)}} = \frac{3(b+a-2s)^2}{4\left((b+a-2s)^2-(b-s)(a-s)\right)}.$$

By rearranging the above equation, we have the following quadratic equation to solve

$$4\alpha s^2 - 4(a+b)\alpha s + \left(1-\frac{4}{3}(1-\alpha)\right)(a+b)^2 + \frac{4}{3}(1-\alpha)ab = 0. \tag{3.28}$$

Finally, by assumption, $s \le a$. So the unique solution s^* to the above quadratic equation is

$$\begin{aligned} s^* &= \frac{\alpha(a+b) - \sqrt{\frac{\alpha}{3}(1-\alpha)(a+b)^2 - \frac{4\alpha}{3}(1-\alpha)(ab)}}{2\alpha} \\ &= \frac{(a+b)}{2} - \frac{(b-a)}{2}\sqrt{\frac{1-\alpha}{3\alpha}}, \end{aligned} \tag{3.29}$$

where $\alpha \in (0, \frac{1}{4}]$.

Indeed, since $s^* \le a$, we get

$$(b-a)\left(1-\sqrt{\frac{1-\alpha}{3\alpha}}\right) \le 0,$$

Since $b-a>0$, we have $\alpha \in (0, \frac{1}{4}]$.

By using (3.29), we obtain

$$\begin{aligned} \pi_{(2,\alpha)}(X) &= \frac{1}{(1-\alpha)}\mathbb{E}[(X-s^*)_+] + s^* \\ &= \frac{1}{2(1-\alpha)}(b+a) - \frac{\alpha}{(1-\alpha)}s^* \\ &= \frac{1}{2}(b+a) + \frac{1}{2}(b-a)\sqrt{\frac{\alpha}{3(1-\alpha)}}, \end{aligned}$$

where $\alpha \in (0, \frac{1}{4}]$.

Note that, if $\alpha = \frac{1}{4}$, then $s^* = a$, and

$$\pi_{(2,\alpha)}(X) = \mathbb{E}[X] + \frac{\sigma(X)}{\sqrt{3}}.$$

Now, let assume $s > a$, then

$$\begin{aligned}
\mathbb{E}[(X-s)_+^p] &= \int_s^b (X-s)^p f(X) dX \\
&= \frac{1}{b-a} \int_s^b (X-s)^p dX \\
&= \frac{(b-s)^{p+1}}{(p+1)(b-a)}.
\end{aligned} \tag{3.30}$$

Similarly, by substituting (3.30) in (3.11) with $p = 1, 2$ and by assumption of $s > a$, the solution s^* is ,

$$s^* = b - \frac{4}{3}(b-a)(1-\alpha), \tag{3.31}$$

where $\alpha \in [\frac{1}{4}, 1)$.

Finally, by plugging (3.31) in (3.12), we obtain

$$\begin{aligned}
\pi_{(2,\alpha)}(X) &= \frac{1}{(1-\alpha)} \mathbb{E}[(X-s^*)_+] + s^* \\
&= b - \frac{4}{9}(b-a)(1-\alpha).
\end{aligned}$$

□

Remark 43. *Note that when $X \sim \mathcal{U}(a, b)$ then*

$$F^{-1}(\alpha) = a + \alpha(b-a), \quad \textit{for any} \;\; \alpha \in (0,1).$$

For any $\alpha \leq \frac{1}{4}$, by Proposition 42, $s^* \leq a$. Also we have

$$s^* = \frac{(a+b)}{2} - \frac{(b-a)}{2}\sqrt{\frac{1-\alpha}{3\alpha}} \leq a < a + \alpha(b-a) = F^{-1}(\alpha).$$

For any $\alpha \geq \frac{1}{4}$, we get

$$s^* - F^{-1}(\alpha) = -\frac{1}{3}(b-a)(1-\alpha) < 0.$$

This confirms the result of Theorem 40.

Moreover, for any $\alpha \leq \frac{1}{4}$, we have

$$s^* < a, \quad \text{and} \quad \pi_{(2,\alpha)}(X) = \mathbb{E}[X] + \sigma(X)\sqrt{\frac{\alpha}{1-\alpha}},$$

which is the same as (3.18) where $s^* < \tilde{y} = a$.

Example 44. *Let $X \sim \mathcal{U}(0,1)$, then using (3.27), we have*

$$\pi_{(2,\alpha)}(X) = \begin{cases} \frac{1}{2} + \frac{1}{2}\sqrt{\frac{\alpha}{3(1-\alpha)}}, & 0 < \alpha \leq \frac{1}{4}, \\ \frac{5}{9} + \frac{4}{9}\alpha, & \frac{1}{4} \leq \alpha \leq 1. \end{cases}$$

Moreover, at $\alpha = \frac{1}{4}$, we have a breaking point, i.e, $s^* = 0$ and

$$\pi_{(2,\alpha)}(X) = \frac{4\mathbb{E}[X_+]}{3} = \frac{2}{3}.$$

In the following, we provide an example that shows in face $s^*_{p,\alpha}(X)$ is not monotone increasing function of X.

Example 45. *If $X \sim \mathcal{U}(0,1)$, then $s^*_{2,\alpha}(X)$ fails to be monotone in X. Indeed for any $\alpha \in (0, \frac{1}{4})$,*

$$s^*_{2,\alpha}(X) = \frac{1}{2} - \frac{1}{2}\sqrt{\frac{1-\alpha}{3\alpha}} < 0.$$

Now, let $Y = 0$, by Theorem 38, $s^*_{2,\alpha}(c) = c$. Thus, $s^*_{p,\alpha}(Y) = 0 > s^*_{p,\alpha}(X)$ while $Y \leq X$.

3.3.2.2 Exponential Distribution

In this subsection, we consider the case with an exponentially distributed risk variable X. The analytical expression for the HG risk measure that obtains in [65], Theorem 7.1 when the Young function is restricted to be a power Young function, is base on assumption that the value $s^*_{p,\alpha} > 0$ (i.e. $s^*_{p,\alpha}$ is inside of support of F). Our study shows that their result holds for all $\alpha \in (1 - \frac{\Gamma(p+1)}{p^p}, 1),\ p \geq 1$. In the following discussion (Proposition 46), the idea of calculation of the $\pi_{(2,\alpha)}(.)$ for the exponential distribution is extended to the case that $\alpha \in (0,1)$ and we show the same result (Theorem 7.1, [65]) holds for any $\alpha > 1 - \frac{\Gamma(p+1)}{p^p}$.

Let us to consider a random variable X such that $X \sim Exp(\lambda)$ where $\lambda > 0$. The Exponential pdf, and cdf are given by,

$$f(x) = \begin{cases} \lambda \exp(-\lambda x), & x \geq 0, \\ 0, & x < 0. \end{cases}, \quad F(x) = \begin{cases} 1 - \exp(-\lambda x), & x \geq 0, \\ 0, & x < 0. \end{cases}$$

With

$$F^{-1}(\alpha) = \frac{-\ln(1-\alpha)}{\lambda}, \quad \text{for any } \alpha \in (0,1). \tag{3.32}$$

Proposition 46. *If $\Phi(x) = x^p,\ p \geq 1$, and $X \sim Exp(\lambda)$, then*

$$\pi_{(p,\alpha)}(X) = \frac{p}{\lambda} - \frac{1}{\lambda} \ln\left(\frac{p^p(1-\alpha)}{\Gamma(p+1)}\right), \quad \textit{for any } \alpha \geq 1 - \frac{\Gamma(p+1)}{p^p}. \tag{3.33}$$

where $0 < s^* < \ln \frac{1}{(1-\alpha)^{\frac{1}{\lambda}}}$.

Moreover, If $p = 2$, then

$$\pi_{(2,\alpha)}(X) = \begin{cases} \frac{1}{\lambda} + \frac{1}{\lambda}\sqrt{\frac{\alpha}{1-\alpha}}, & \alpha \leq \frac{1}{2}, \\ \frac{2}{\lambda} - \frac{1}{\lambda} \ln(2(1-\alpha)), & \alpha \geq \frac{1}{2}. \end{cases} \tag{3.34}$$

Proof. Let assume $s \geq 0$, we evaluate $\mathbb{E}[(X-s)_+^p]$ as follows.

$$\begin{aligned}
\mathbb{E}[(X-s)_+^p] &= \int_s^\infty (X-s)^p f(X) dX \\
&= \lambda exp(-\lambda s) \int_0^\infty X^p exp(-\lambda X) dX \\
&= \frac{\Gamma(p+1)}{\lambda^p} exp(-\lambda s) \\
&= \frac{p\Gamma(p)}{\lambda^p} exp(-\lambda s), \qquad (3.35)
\end{aligned}$$

where $\Gamma(p) = \int_0^\infty x^{p-1} exp(-x) dx$.

Then, by substituting (3.35) in (3.11), we have the following equation.

$$\frac{\left(\mathbb{E}\left[(X-s)_+^{p-1}\right]\right)^p}{\left(\mathbb{E}\left[(X-s)_+^p\right]\right)^{p-1}} = \frac{\Gamma(p)\exp(-\lambda s)}{p^{p-1}} = 1-\alpha.$$

Then the unique positive solution s^* is

$$s^* = \frac{-1}{\lambda} \ln\left(\frac{p^{p-1}(1-\alpha)}{\Gamma(p)}\right) = \frac{-1}{\lambda} \ln\left(\frac{p^p(1-\alpha)}{\Gamma(p+1)}\right). \qquad (3.36)$$

where $\alpha > 1 - \frac{\Gamma(p+1)}{p^p}$.

Finally, by using (3.35), and (3.36), we get

$$\begin{aligned}
\pi_{(p,\alpha)}(X) &= \frac{1}{(1-\alpha)^{\frac{1}{p}}} \left(\mathbb{E}\left[(X-s^*)_+^p\right]\right)^{\frac{1}{p}} + s^* \\
&= \frac{p}{\lambda} - \frac{1}{\lambda} \ln\left(\frac{p^p(1-\alpha)}{\Gamma(p+1)}\right).
\end{aligned}$$

Moreover, if $p = 2$, by Theorem 38 (iv), for any $\alpha \geq \frac{1}{2}$, we have $s^* \geq 0$. Thus, we get

$$\pi_{(2,\alpha)}(X) = \frac{2}{\lambda} - \frac{1}{\lambda} \ln(2(1-\alpha)).$$

Now let p be an integer, then by Theorem 38, $s^* < 0$, for any $\alpha \leq \frac{1}{2}$, and we have

$$\begin{aligned}
\mathbb{E}[(X-s)_+^p] &= \int_s^{\infty} (X-s)^p f(X) dX \\
&= \lambda \int_0^{\infty} (X-s)^p exp(-\lambda X) dX \\
&= \lambda \sum_{k=0}^{p} \binom{p}{k} (-s)^k \int_0^{\infty} X^{p-k} exp(-\lambda X) dX \\
&= \sum_{k=0}^{p} \binom{p}{k} \frac{1}{\lambda^{p-k}} (-s)^k \Gamma(p-k+1) \\
&= \sum_{k=0}^{p} \frac{1}{\lambda^{p-k}} \frac{p!}{k!} (-s)^k.
\end{aligned}$$

When $p = 2$, the optimizer s^* is the solution to the following equation

$$\frac{(\frac{1}{\lambda} - s)^2}{\frac{1}{\lambda^2} + (\frac{1}{\lambda} - s)^2} = 1 - \alpha.$$

Since, $s^* < 0$, so we have

$$s^* = \frac{1}{\lambda} - \sqrt{\frac{1-\alpha}{\alpha}} \frac{1}{\lambda}.$$

and

$$\begin{aligned}
\pi_{(2,\alpha)}(X) &= \frac{1}{\lambda\sqrt{\alpha(1-\alpha)}} + \frac{1}{\lambda} - \sqrt{\frac{1-\alpha}{\alpha}} \frac{1}{\lambda} \\
&= \frac{1}{\lambda} + \frac{1}{\lambda} \sqrt{\frac{\alpha}{1-\alpha}}.
\end{aligned}$$

□

Remark 47. *Note that when $X \sim Exp(\lambda)$, then*

$$F^{-1}(\alpha) = -\frac{1}{\lambda} \ln(1-\alpha).$$

For any $\alpha \geq \frac{1}{2}$, the optimizer s^* is given by

$$s^* = \frac{-1}{\lambda} \ln (2 - 2\alpha).$$

Therefore,

$$s^* - F^{-1}(\alpha) = -\frac{1}{\lambda}(\ln 2(1-\alpha) - \ln (1-\alpha) = -\frac{\ln 2}{\lambda} < 0.$$

When $\alpha \leq \frac{1}{2}$, then $s^* \leq 0$, and

$$s^* = \frac{1}{\lambda} - \sqrt{\frac{1-\alpha}{\alpha}}\frac{1}{\lambda} < 0 \leq F^{-1}(\alpha).$$

Thus, the result is consistent with Theorem 40. Also note that

$$\pi_{(2,\alpha)}(X) = \mathbb{E}[X] + \sigma(X)\sqrt{\frac{\alpha}{1-\alpha}}, \quad \textit{for any } \alpha < \frac{1}{2}, \tag{3.37}$$

which confirms the result of Proposition 40, the equation (3.18).

Example 48. *Let $X \sim Exp(1)$, then*

$$F^{-1}(\alpha) = -\ln(1-\alpha), \quad \textit{for any } \alpha \in (0,1).$$

Using (3.35) when $p = 1$ and (3.34), we get

$$\begin{aligned} AV@R_\alpha &= \frac{\mathbb{E}[(X + \ln(1-\alpha))_+]}{1-\alpha} - \ln(1-\alpha) \\ &= 1 - \ln(1-\alpha), \end{aligned}$$

and

$$\pi_{(2,\alpha)}(X) = \begin{cases} 1 + \sqrt{\frac{\alpha}{1-\alpha}}, & \alpha \leq \frac{1}{2}, \\ 2 - \ln\left(2\left(1-\alpha\right)\right), & \alpha \geq \frac{1}{2}. \end{cases} \tag{3.38}$$

Note that at $\alpha = \frac{1}{2}$*,* $s^* = 0$ *and* $\pi_{(2,\alpha)}(X) = 2$*.*

3.3.2.3 Normal Distribution

For the final step of our study, we picked the standard normal distribution; unfortunately, we were unable to find the explicit solution to the equation (3.11). Therefore, we employ a numerical procedure to solve the equation (3.11).

In the following, suppose that the return of a financial asset X is normally distributed, $X \sim \mathcal{N}(0,1)$, where $\mathcal{N}(0,1)$ is a standard normal random variable, where pdf, cdf are given by

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right), \quad \Phi(x) = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{x}{\sqrt{2}}\right)\right),$$

where erf(.), commonly known as error function, is

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp\left(-t^2\right) dt.$$

Remark 49. *Recall that* $X \sim \mathcal{N}(\mu, \sigma)$*, then*

$$\mathrm{AV@R}_\alpha(X) = \mu + \sigma \frac{\varphi(\Phi^{-1}(\alpha))}{1-\alpha}. \tag{3.39}$$

In particular, when $X \sim \mathcal{N}(0,1)$*,*

$$\mathrm{AV@R}_\alpha(X) = \frac{\varphi(\Phi^{-1}(\alpha))}{1-\alpha}.$$

Proposition 50. *Let* $X \sim \mathcal{N}(0,1)$*, for any* $\alpha \in (0,1)$*, then*

$$\pi_{(2,\alpha)}(X) = s^* + \frac{\varphi(s^*)}{1-\alpha}\left(1 - s^* m(s^*)\right), \tag{3.40}$$

where $s^* < \Phi^{-1}(\alpha)$ *and it is the unique solution to the nonlinear equation*

$$\xi(s) = \varphi(s)\left(1 - s.m(s)\right)^2 - (1-\alpha)\left((s^2+1)m(s) - s\right), \tag{3.41}$$

where $m(s) := \frac{\Phi(-s)}{\varphi(-s)}$ *is the mills ratio for the standard normal distribution.*

Proof. First, we evaluate $E[(X-s)_+]$, $E[(X-s)_+^2]$

$$\begin{aligned}
\mathbb{E}[(X-s)_+] &= \int_s^\infty (X-s)\varphi(X)dX \\
&= -\frac{s\left(\operatorname{erf}\left(\frac{s}{\sqrt{2}}\right)-1\right)}{2} + \frac{e^{-\frac{s^2}{2}}}{\sqrt{2\pi}} \\
&= \varphi(s) - s(1-\Phi(s)) = \varphi(s) - s\Phi(-s). \qquad (3.42)
\end{aligned}$$

Similarly, by using integrating by part, we get

$$\begin{aligned}
\mathbb{E}[(X-s)_+^2] &= \int_s^\infty (X-s)^2\varphi(X)dX \\
&= -\frac{\left(s^2+1\right)\left(\operatorname{erf}\left(\frac{s}{\sqrt{2}}\right)-1\right)}{2} - s\frac{e^{-\frac{s^2}{2}}}{\sqrt{2\pi}} \\
&= \left(s^2+1\right)(1-\Phi(s)) - s\varphi(s) = \left(s^2+1\right)\Phi(-s) - s\varphi(s). \qquad (3.43)
\end{aligned}$$

By applying (3.42) and (3.43) into the equation (3.11), we obtain

$$\frac{(\varphi(s)-s\Phi(-s))^2}{(s^2+1)\,\Phi(-s)-s\varphi(s)} = 1-\alpha.$$

Then by Theorem 40, $s^* < \Phi^{-1}(\alpha)$ and s^* is the unique solution to the following equation,

$$(\varphi(s)-s\Phi(-s))^2 - (1-\alpha)\left(\Phi(-s) - s\left(\varphi(s)-s\Phi(-s)\right)\right) = 0$$

Assume $\varphi(s) \neq 0$. Therefore, we have the following equation to solve.

$$\xi(s) = \varphi(s)\left(1-s.m(s)\right)^2 - (1-\alpha)\left((s^2+1)m(s)-s\right),$$

where $m(s) := \frac{\Phi(-s)}{\varphi(s)}$.

Using (3.11), we obtain

$$[\Phi(-s^*) - s^*(\varphi(s^*) - s^*\Phi(-s^*))]^{\frac{1}{2}} = \frac{1}{\sqrt{1-\alpha}}[\varphi(s^*) - s^*\Phi(-s^*)],$$

where s^* is the unique solution of the function $\xi(s)$,

Finally, we get

$$\begin{aligned}\pi_{(2,\alpha)}(X) &= \frac{1}{\sqrt{1-\alpha}}\sqrt{\mathbb{E}[(s^*-X)_+^2]} + s^* \\ &= \frac{1}{1-\alpha}[\varphi(s^*) - s^*\Phi(-s^*)] + s^* \\ &= s^*\left(1 - \frac{1}{1-\alpha}\Phi(-s^*)\right) + \frac{1}{1-\alpha}\varphi(s^*).\end{aligned}$$

□

Corollary 51. *Let X be random variable such that $X \sim \mathcal{N}(0,1)$ with $\pi_{(2,\alpha)}$. Then*

(i) At $\alpha = 1 - \frac{1}{\pi}$, the optimizer is $s^* := s_{2,\alpha} = 0$ and $\pi_{(2,\alpha)}(X) = \sqrt{\frac{\pi}{2}}$.

(ii) For any $\alpha > 1 - \frac{1}{\pi}$, s^* is positive and

$$0 < s^* \leq \pi_{(2,\alpha)}(X) \leq \frac{1}{\sqrt{2\pi}(1-\alpha)} + \Phi^{-1}(\alpha). \tag{3.44}$$

(iii) For any $\alpha < 1 - \frac{1}{\pi}$, s^* is negative and

$$\frac{\alpha}{1-\alpha}s^* \leq \pi_{(2,\alpha)}(X) \leq \frac{1}{\sqrt{2\pi}(1-\alpha)} - \frac{\alpha}{1-\alpha}s^*. \tag{3.45}$$

Proof. (i) It is easy to check that at $\alpha = 1 - \frac{1}{\pi}$, $s^* = 0$. Indeed, by letting $s^* = 0$, and using (3.41), we get

$$\frac{1}{\sqrt{2\pi}} - (1-\alpha)m(0) = 0.$$

So solving the above equation for α, we have

$$\begin{aligned}\alpha &= 1 - \frac{1}{m(0)\sqrt{2\pi}} \\ &= 1 - \frac{1}{\pi}.\end{aligned}$$

Then

$$\pi_{(\alpha,2)}(X) = \frac{1}{1-\alpha}\varphi(0) = \sqrt{\frac{\pi}{2}}.$$

(ii) For any $\alpha > 1 - \frac{1}{\pi}$, by Theorem 38, $s^* > 0$. Recall that the mills ratio $m(s)$ for the standard normal distribution has the following bounds for all $s > 0$,

$$\frac{s}{s^2+1} \leq m(s) \leq \frac{1}{s}. \tag{3.46}$$

Using (3.46), we get

$$0 \leq \frac{\varphi(s^*)}{1-\alpha}\left(1 - s^*.m(s^*)\right) \leq \frac{\varphi(s^*)}{({s^*}^2+1)(1-\alpha)}.$$

Therefore, $0 < s^* \leq \Phi^{-1}(\alpha)$ for any $\alpha \in (1 - \frac{1}{\pi}, 1)$, and

$$0 < s^* \leq \pi_{(2,\alpha)}(X) \leq \frac{\varphi(s^*)}{({s^*}^2+1)(1-\alpha)} + s^* \leq \frac{1}{\sqrt{2\pi}(1-\alpha)} + \Phi^{-1}(\alpha).$$

(iii) For any $\alpha < 1 - \frac{1}{\pi}$, $s^* < 0$ and using (3.46) again for $s < 0$, we have

$$\frac{-s^*}{\varphi(s^*)} \leq 1 - s^* m(s^*) \leq \frac{1}{{s^*}^2+1} - \frac{s^*}{\varphi(s^*)}.$$

Then,

$$\frac{-s^*}{1-\alpha} \leq \frac{\varphi(s^*)}{1-\alpha}\left(1 - s^* m(s^*)\right) \leq \frac{\varphi(s^*)}{(1-\alpha)({s^*}^2+1)} - \frac{s^*}{1-\alpha}.$$

And finally

$$
\begin{aligned}
-s^* \frac{\alpha}{1-\alpha} \leq \pi_{(2,\alpha)}(X) &\leq \frac{\varphi(s^*)}{(1-\alpha)(s^{*2}+1)} - \frac{s^*}{1-\alpha} + s^* \\
&\leq \frac{1}{\sqrt{2\pi}(1-\alpha)} - s^* \frac{\alpha}{1-\alpha}.
\end{aligned}
$$

$\square$

3.4 An Empirical Study

This section concentrates on an estimation method to quantify the $\pi_{(2,\alpha)}(.)$ when the data are assumed to come from a non-particular distribution. We use the empirical method, known as the Historical simulation, to approximate the $\pi_{(2,\alpha)}(.)$. Recall that the empirical method is a non-parametric method that needs no assumptions regarding the data or distribution. Being relatively easy to implement and not computationally intensive makes this method the most extensively used method in the financial industry and academic studies.

3.4.1 Empirical Distribution and Empirical Quantile

The Historical simulation method is based on the construction of the empirical distribution function using historical data. In what follows, we let $X_1, \dots, X_n$ be independent and identically distributed (i.i.d.) random variables with (unknown) distribution function F.

Definition 52. *Given $X_1, \dots, X_n$, the order statistics $X_{(1)}, \dots, X_{(n)}$ is defined as the permutation of $X_1, \dots, Y_n$ such that $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$.*

The empirical distribution F_n is constructed by placing at each observation X_i a mass $1/n$. Thus F_n may be represented as

$$F_n(x) := \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{[X_k \leq x]}. \tag{3.47}$$

The empirical quantile function F_n^{-1} is the quantile function of the empirical distribution function F_n define on the interval $(0, 1)$ given by

$$F_n^{-1}(\alpha) := \inf \{t \ : F_n(t) \geq \alpha\} = X_{(k)}, \tag{3.48}$$

where k is chosen such that $\frac{k-1}{n} < \alpha \leq \frac{k}{n}$.

Recall that given sample $X_1, \dots, X_n$ of independent copies of X, the empirical estimate of $\mathrm{V@R}_\alpha(X)$ can be expressed using order statistics. Note that $F_n(X_{(k)}) = \frac{k}{n}$ and

$$\begin{aligned}\widehat{\text{V@R}}_\alpha(X) &:= \inf\left\{X_{(k)} : \frac{k}{n} \geq \alpha\right\} \\ &= \inf\left\{X_{(k)} : k \geq n\alpha\right\} \\ &= X_{(\lceil n\alpha\rceil)},\end{aligned}$$

where $\lceil y\rceil := \inf\{k \in \mathbb{Z} : k \geq y\}$ which is the ceiling function that gives the smallest integer not less than y.

3.4.2 Empirical Estimation: The Average Value-at-risk

Here we recall the algorithm of approximating the $AV@R_\alpha(.)$ in the form of the optimization problem . Recall that,

$$\text{AV@R}_\alpha(X) = s^* + \frac{\mathbb{E}[(X - s^*)_+]}{1-\alpha}, \tag{3.49}$$

where $s^* = F^{-1}(\alpha)$.

The integral in (3.49) can be calculated approximately or exactly. For example, in the case of a discrete sample $(X_i)_{1\leq i\leq n}$, the rectangles method will lead to the following formula:

$$\widehat{\text{AV@R}}_\alpha(X) = s^* + \frac{\sum_{i=1}^n (X_i - s^*)_+}{n(1-\alpha)}, \tag{3.50}$$

which is the sample version of the equation (3.49).

Proposition 53. *For a given $\alpha \in (0,1)$, suppose that $\lceil \alpha n\rceil \geq 2$. Let $X_1, \ldots, X_n$ be i.i.d. random variables, then*

$$\widehat{\text{AV@R}}_\alpha(X) = \sum_{i=\lfloor \alpha n\rfloor+1}^{n} \frac{X_{(i)}}{n(1-\alpha)} + \left(1 - \frac{n - \lfloor \alpha n\rceil}{n(1-\alpha)}\right) X_{(\lceil \alpha n\rfloor)}. \tag{3.51}$$

where $(X_{(i)})_{1\leq i\leq n}$ is order statistics.

Moreover, If $\lceil \alpha n \rceil$ is an integer then

$$\widehat{\text{AV@R}}_\alpha(X) = \sum_{i=n\alpha+1}^{n} \frac{X_{(i)}}{n(1-\alpha)}.$$

Proof. For a given $\alpha \in (0,1)$, since $\lceil \alpha n \rceil \geq 2$, there exist k such that $\alpha \in (\frac{k-1}{n}, \frac{k}{n}]$. Since $s^* = F^{-1}(\alpha)$, so $X_{(k-1)} < X_{(k)} = s^*$, where $(X_{(i)})_{1 \leq i \leq n}$ is order statistics.

By applying $s^* = F^{-1}(\alpha)$ in (3.50), we have

$$\begin{aligned}
\widehat{\text{AV@R}}_\alpha(X) &= \frac{\sum_{i=k+1}^{n}(X_{(i)} - s^*)}{n(1-\alpha)} + s^* \\
&= \frac{1}{n(1-\alpha)} \sum_{i=\lceil \alpha n \rceil + 1}^{n} (X_{(i)} - s^*) + s^* \\
&= \sum_{i=\lceil \alpha n \rceil + 1}^{n} \frac{X_{(i)}}{n(1-\alpha)} + \left(1 - \frac{n - \lceil \alpha n \rceil}{n(1-\alpha)}\right) X_{(\lceil \alpha n \rceil)}.
\end{aligned}$$

If $\lceil \alpha n \rceil$ is an integer, then the expression in the last display reduces to the sample mean of the $\lceil \alpha n \rceil$ largest losses. Therefore

$$\widehat{\text{AV@R}}_\alpha(X) = \sum_{i=n\alpha+1}^{n} \frac{X_{(i)}}{n(1-\alpha)}.$$

□

This gives a similar formula that derived by Peracchi and Tanase in [58] which is a linear combination of extreme order statistics and change with the sample size.

Algorithm 1: The Historical Simulation for $\text{AV@R}_\alpha(.)$:

Using above, we will have the following algorithm to calculate $\widehat{\text{AV@R}}_\alpha(.)$.

Step 1: Take a large number of observations $(X_i)_i$, $i = 1, ..., n$,

Step 2: Sort $(X_i)_{1 \leq i \leq n}$ as their order statistics.

Step 3: Under the level of $\alpha \in (0,1)$, find $\lceil \alpha n \rceil$.

Step 4: Calculate $\widehat{\mathrm{AV@R}}_\alpha(.)$ as

$$\sum_{i=\lceil \alpha n\rceil+1}^{n} \frac{X_{(i)}}{n(1-\alpha)} + \left(1 - \frac{n-\lceil \alpha n\rceil}{n(1-\alpha)}\right) X_{(\lceil \alpha n\rceil)}.$$

3.4.3 Empirical Estimation: The Haezendonck-Goovaerts Risk Measure

In this section, we describe our propose method for constructing the empirical estimation of the HG risk measure $\pi_{(2,\alpha)}(.)$. The empirical implementation of the HG risk measure $\pi_{(2,\alpha)}(.)$ can be done by combining a similar technique as we describe in Section 3.4.2 and Theorem 36. As we assume in section 3.4.2, let $X_1, \ldots, X_n$ be a sequence of i.i.d random variables. Also, assume $F(.)$ be their common distribution function. Note that by the definition of the HG risk measure, we could define $\pi_{(2,\alpha)}(F)$ to be equal $\pi_{(2,\alpha)}(X)$. With $F_n(.)$ be the empirical distribution function, it is easy to see that $\pi_{(2,\alpha)}(F_n)$ is a well-defined random variable. Then by [3] Theorem 1, $\pi_{(2,\alpha)}(F_n)$ is a natural non-parametric estimators for $\pi_{(2,\alpha)}(F)$, i.e.

$$\lim_{n\to\infty} \pi_{(2,\alpha)}(F_n) = \pi_{(2,\alpha)}(F),$$

almost surely.

Note that, Theorem 36 already gives an algorithm on how to empirically calculate the $\pi_{(2,\alpha)}(.)$. By Theorem 36, a natural candidate to approximating the value of s^*, for any $\alpha < \frac{n-1}{n}$ would be the unique solution to the equation (3.9) where $s^* \in (-\infty, X_{(n-1)})$. The following proposition shows how one can implement the value of $\widehat{\pi}_{(2,\alpha)}(.)$ empirically.

Proposition 54. *For given $\alpha \in (0,1)$, suppose that $\lceil n\alpha\rceil \geq 2$. Let $X_1, \ldots, X_n$ be i.i.d. random variables and X that follows the corresponding empirical distribution function F_n. Then if $\alpha < \frac{n-1}{n}$, the minimize $s^* := s^*_{2,\alpha} \in (-\infty, X_{(\lceil n\alpha\rceil)})$. Moreover, let k be the largest index such that for any $m \leq k$, $X_{(m)} \leq s^*$, then*

$$\widehat{\pi}_{(2,\alpha)}(X) = \bar{X}_{(k)} + \left(\frac{k - n(1-\alpha)}{n(1-\alpha)}\right)^{\frac{1}{2}} \sigma_{(k)}. \tag{3.52}$$

where $\bar{X}_{(k)}$, $\sigma^2_{(k)}$ are the truncated sample mean and variance respectively, i.e.

$$\bar{X}_{(k)} := \frac{1}{(n-k)} \sum_{i=k+1}^{n} X_{(i)}, \quad \sigma^2_{(k)} := \frac{1}{(n-k)} \sum_{i=k+1}^{n} X^2_{(i)} - \bar{X}^2_{(k)}.$$

Proof. Pick any $\alpha \in (0,1)$, since $\lceil n\alpha \rceil \geq 2$ so there exist i, such that $\alpha \in (\frac{i-1}{n}, \frac{i}{n}]$ and $F_n^{-1}(\alpha) = X_{(\lceil n\alpha \rceil)}$.

Let s^* be a minimizer for the equation (3.11) and let $k \in \{1, \ldots, n-1\}$ be such that for any $m \leq k$, $X_{(m)} \leq s^*$, so $s^* < X_{(k+1)}$. If $k = n-1$, then $s^* \in [X_{(n-1)}, X_{(n)})$ and $\widehat{\pi}_{(2,\alpha)} = X_{(n)}$. For any $k < n-1$, with a similar process as Theorem 40, it is easy to show that $k < \lceil n\alpha \rceil$. Indeed, under conditions of Theorem 36, $s^* \in (-\infty, X_{(n-1)})$ is the unique solution to the following equation

$$\begin{aligned}
\frac{(\mathbb{E}[(X-s^*)_+])^2}{\mathbb{E}[(X-s^*)_+^2]} &= \frac{\left(\frac{1}{n}\sum_{i=1}^{n}(X_{(i)}-s^*)_+\right)^2}{\frac{1}{n}\sum_{i=1}^{n}\left((X_{(i)}-s^*)_+\right)^2} \\
&= \frac{\left(\sum_{i=k+1}^{n}(X_{(i)}-s^*)\right)^2}{n\sum_{i=k+1}^{n}\left(X_{(i)}-s^*\right)^2} = 1-\alpha.
\end{aligned}$$

For simplification, let

$$\sum_{i=k+1}^{n} X_{(i)} = (n-k)\bar{X}_{(k)}, \quad \text{and} \quad \sum_{i=k+1}^{n} X^2_{(i)} = (n-k)\left(\sigma^2_{(k)} + \bar{X}^2_{(k)}\right).$$

Then

$$\begin{aligned}
\frac{\left(\sum_{i=k+1}^{n}(X_{(i)}-s^*)\right)^2}{\sum_{i=k+1}^{n}(X_{(i)}-s^*)^2} &= \frac{(n-k)^2(\bar{X}_{(k)}-s^*)^2}{(n-k)\left(s^{*2}+\bar{X}^2_{(k)}-2s^*\bar{X}_{(k)}+\sigma^2_{(k)}\right)} \\
&= \frac{(n-k)(\bar{X}_{(k)}-s^*)^2}{\left(\bar{X}_{(k)}-s^*\right)^2+\sigma^2_{(k)}} = n(1-\alpha).
\end{aligned}$$

Note that $s^* \leq \bar{X}_{(k)}$. By rearrangement of the above equation, we have

$$(n\alpha - k)(\bar{X}_{(k)} - s^*)^2 = n(1-\alpha)\sigma^2_{(k)}. \tag{3.53}$$

The equation (3.53) has a solution if $k < \lceil n\alpha \rceil$, therefore, $s^* < X_{(\lceil n\alpha \rceil)}$.

Solving (3.53) for s^*, we obtain

$$s^* = \bar{X}_{(k)} - \left(\frac{n(1-\alpha)}{n\alpha - k}\right)^{\frac{1}{2}} \sigma_{(k)}. \tag{3.54}$$

So, using (3.54), and $\left(\sum_{i=k+1}^{n}(X_{(i)} - s^*)\right)^2 = n(1-\alpha)\left(\sum_{i=k+1}^{n}\left(X_{(i)} - s^*\right)^2\right)$, we have

$$\begin{aligned}
\widehat{\pi}_{(2,\alpha)}(X) &= \frac{1}{\sqrt{n(1-\alpha)}}\left(\sum_{i=k+1}^{n}\left(X_{(i)} - s^*\right)^2\right)^{\frac{1}{2}} + s^* \\
&= \frac{1}{n(1-\alpha)}\sum_{i=k+1}^{n}\left(X_{(i)} - s^*\right) + s^* \\
&= \frac{n-k}{n(1-\alpha)}\bar{X}_{(k)} - \frac{n\alpha - k}{n(1-\alpha)}\left(\bar{X}_{(k)} - \left(\frac{n(1-\alpha)}{n\alpha - k}\right)^{\frac{1}{2}}\sigma_{(k)}\right) \\
&= \bar{X}_{(k)} + \left(\frac{n\alpha - k}{n(1-\alpha)}\right)^{\frac{1}{2}}\sigma_{(k)}.
\end{aligned}$$

Finally, we get

$$\widehat{\pi}_{(\alpha,2)}(X) = \bar{X}_{(k)} + \left(\frac{n\alpha - k}{n(1-\alpha)}\right)^{\frac{1}{2}}\sigma_{(k)}. \tag{3.55}$$

□

Remark 55. *In addition to the above Proposition, it is worthwhile mentioning that there exist $\tilde{\alpha} \in (0,1)$, such that $s^*_{\alpha} \leq X_{(1)}$ for any $\alpha \leq \tilde{\alpha}$. Indeed, by letting $s^* = X_{(1)}$ in the equation (3.53), at $\tilde{\alpha}$, we get*

$$\tilde{\alpha} = \frac{\sigma^2(X)}{(\bar{X} - X_{(1)})^2 + \sigma^2(X)}.$$

Therefore, for any $\alpha \leq \tilde{\alpha}$, $s^* \leq X_{(1)}$ and

$$\widehat{\pi}_{(2,\alpha)}(X) = \bar{X} + \sigma(X)\sqrt{\frac{\alpha}{1-\alpha}}, \quad s^* = \bar{X} - \sigma(X)\sqrt{\frac{1-\alpha}{\alpha}}.$$

Example 56. *Let $\Omega = \{w_1, w_2, w_3\}$, with $\mathbb{P}(\{w_i\}) = \frac{1}{3}$, for $i = 1, 2, 3$. Now consider the following random variable*

$$X := \begin{cases} 2, & \textit{on } w_1, \\ 4, & \textit{on } w_2, \\ 6, & \textit{on } w_3 \end{cases}$$

Then truncated sample mean and variance are $\bar{X} = \{4, 5, 6\}$*,* $\sigma^2 = \{\frac{8}{3}, 1, 0\}$ *respectively. Using Proposition 54, we obtain*

$$s^* = \begin{cases} 4 - \sqrt{\frac{8(1-\alpha)}{3\alpha}}, & \alpha \leq \frac{2}{5}, \\ 5 - \sqrt{\frac{3(1-\alpha)}{3\alpha-1}}, & \frac{2}{5} < \alpha \leq \frac{2}{3}, \\ 6, & \frac{2}{3} < \alpha < 1. \end{cases} \quad \widehat{\pi}_{(2,\alpha)}(X) = \begin{cases} 4 + \sqrt{\frac{3\alpha}{8(1-\alpha)}}, & \alpha \leq \frac{2}{5}, \\ 5 + \sqrt{\frac{3\alpha-1}{3(1-\alpha)}}, & \frac{2}{5} < \alpha \leq \frac{2}{3}, \\ 6, & \frac{2}{3} < \alpha < 1. \end{cases}$$

Note that, actually for any $\alpha < \frac{2}{5}$*,* $s^* < X_{(1)}$.

Algorithm 2: The Historical Simulation for $\pi_{(2,\alpha)}(.)$:

Using the above, we have the following algorithm to estimate $\pi_{(2,\alpha)}(.)$

Step 1: Take a large number of observations $(X_i)_i, \ i = 1, \ldots, n$ of random variable X

Step 2: Sort $(X_i)_{1 \leq i \leq n}$ as their order statistics.

Step 3: Pick a risk tolerance level $\alpha \in (0, 1)$, and let $k = \lceil n\alpha \rceil$. Observe that, we only consider the outcomes $\left\{X_{(\lceil n\alpha \rceil + 1)}, \ldots, X_{(n)}\right\}$.

Step 4: Find the empirical truncating mean $\bar{X}_{(k)}$ and variance $\sigma^2_{(k)}$.

$$\bar{X}_{(k)} := \frac{1}{(n-k)} \sum_{i=k+1}^{n} X_{(i)}, \quad \sigma^2_{(k)} := \frac{1}{(n-k)} \sum_{i=k+1}^{n} X^2_{(i)} - \bar{X}^2_{(k)}.$$

Step 5: Calculate

$$s^*_{test} = \bar{X}_{(k)} - \left(\frac{n(1-\alpha)}{n\alpha - k}\right)^{\frac{1}{2}} \sigma_{(k)}.$$

Step 6: If $X_{(k-1)} < s^*_{test} \leq X_{(k)}$ then $s^* = s_{test}$ and go to Step 8. If not, replace k by $k-1$,

Step 7: If $k = 0$ go to Step 8 otherwise go to Step 4.

Step 8: Calculate $\widehat{\pi}_{(2,\alpha)}$ as

$$\bar{X}_{(k)} + \left(\frac{n\alpha - k}{n(1-\alpha)} \right)^{\frac{1}{2}} \sigma_{(k)}. \tag{3.56}$$

Proposition 57. *Let $X_1, \ldots, X_n$ be i.i.d. random variables and $X \in \mathcal{Y}_2$. Then the optimizer $s^*_{2,\alpha}$ for the equation (3.50) is increasing as α is increasing.*

Proof. Let assume $s^*_1 := s^*_{2,\alpha_1}$ and $s^*_2 := s^*_{2,\alpha_2}$ be solutions to the equation (3.50) at α_1 and α_2 respectively where $\alpha_1 < \alpha_2$. Then using (3.53), we have

$$\left(s^*_i - \bar{X}_{(k_i)}\right)^2 = \frac{\sigma^2_{(k_i)}}{\beta_i}, \quad \text{for } i = 1, 2. \tag{3.57}$$

where $\beta_i = (\frac{n-k_i}{n(1-\alpha_i)} - 1)$.

Let assume that $s_1 > s_2$, so $k_1 > k_2$ and $\sigma_{(k_1)} > \sigma_{(k_2)}$. Note that $\bar{X}_{(k_i)} - s^*_i \geq 0$ for $i = 1, 2$ and

$$\left(\bar{X}_{k_2} - s^*_2\right)^2 > \left(\bar{X}_{k_1} - s^*_1\right)^2.$$

Using (3.57), we get

$$0 > \frac{1}{\beta_1}\sigma^2_{k_1} - \frac{1}{\beta_2}\sigma^2_{k_2} > \left(\frac{1}{\beta_1} - \frac{1}{\beta_2}\right)\sigma^2_{k_2}.$$

Then

$$\frac{n-k_2}{n(1-\alpha_2)} - 1 = \beta_2 < \beta_1 = \frac{n-k_1}{n(1-\alpha_1)} - 1.$$

Finally, since $k_2 < k_1$, we get

$$\frac{n-k_2}{n(1-\alpha_2)} < \frac{n-k_1}{n(1-\alpha_1)} < \frac{n-k_2}{n(1-\alpha_1)}.$$

Therefore,

$$n(1-\alpha_1) < n(1-\alpha_2)$$

which is contradiction by the assumption of $\alpha_1 < \alpha_2$. So $s^*_{2,\alpha}$ is increasing as the α level is increasing. $\square$

Remark 58. *In general, when $\Phi(x) = x^p$, $p > 2$, for any given $\alpha \in (0,1)$, we have*

$$\widehat{\pi}_{(p,\alpha)}(X) = s^* + \left(\frac{\sum_{i=1}^{n} \left((X_{(i)} - s^*)_+ \right)^p}{n(1-\alpha)} \right)^{\frac{1}{p}}. \tag{3.58}$$

where s^* is the solution to the equation

$$\left(\sum_{i=1}^{n} \left((X_{(i)} - s)_+ \right)^{p-1} \right)^p - n(1-\alpha) \left(\sum_{i=1}^{n} \left((X_{(i)} - s)_+ \right)^p \right)^{p-1} = 0. \tag{3.59}$$

In a general format, implementation of such a risk measure can be done via, for example, cone programming. Also, one can find the solution to the equation (3.59) numerically by standard optimization packages such as nleqslv, BB *packages in RStudio and plugging in the equation (3.58).*

In the next chapter, we illustrate a practical example of an application of the HG risk measure on real data using the historical simulation method base on the assumption of data continuously. We show that this family of risk measures can be a promising tool in risk management.

Chapter 4

The Haezendonck-Goovaerts Risk Measure: Technical Report

In this Chapter, we conduct numerical studies based on methods and results of the previous Chapter. In addition, in Section 4.2, as a case study, we investigate the behaviour of the HG risk measure $\pi_{(p,\alpha)}(.), p \geq 1$ by employing the Historical Simulation method for real data using different scenarios and periods. Note that the focus in this study is not the analysis of issues of modelling or even details on different financial applications, but the behaviour of $\pi_{(p,\alpha)}(.)$ when applied to financial data.

4.1 Technical Report

In this Section, we apply the methods described in Sections 3.4.2 and 3.4.3, to obtain the value of $\pi_{(p,\alpha)}(.)$ numerically using uniform, exponential, and normal distributions. We expose some plots based on the method that we describe to visualize the optimal value $\pi_{(p,\alpha)}(.), p \geq 1$ and its optimizer $s^*_{p,\alpha}$.

4.1.1 Software

The code for this thesis is written in RStudio, and can be found here[1]. For conceptual treatment of distributions, the package *distrEx* ([46]) and *RobExtremes*([42]) are used. The package *pracma* ([13]) is used for the numerical integration and the numerical optimization purposes of this thesis. Finally, for plotting, *ggplot2* ([68]) time series analysis, *zoo* ([69]) and for computing expected shortfall, *cvar* ([14]) have been

[1] https://github.com/arahsepar/HG-Risk-Measures.git

used. All Computations are done with a personal laptop with *Intel(R), Core(TM) i7-8565U CPU @ 1.80GHz 1.99 GHz, 64-based processor*.

4.1.2 Error Estimation

Using the results of Chapter 3, one can model an approximation method to calculate $\pi_{(2,\alpha)}(.)$. With fixed α, the first step is to estimate the value $\widetilde{s}^* := \widetilde{s}^*_{2,\alpha}$ as the unique root of the function ξ where

$$\xi(s) := (1-\alpha)^{\frac{1}{2}} \left(\mathbb{E}[(X-s)_+^2]\right)^{\frac{1}{2}} - \mathbb{E}[(X-s)_+]. \tag{4.1}$$

Using (4.1), one can obtain

$$\widetilde{\pi}_{(p,\alpha)}(X) = \frac{1}{1-\alpha}\xi(\tilde{s}^*) + \frac{1}{1-\alpha}\mathbb{E}[(X-\tilde{s}^*)_+] + \tilde{s}^*. \tag{4.2}$$

Note that, when we able to obtain the exact unique root of the function ξ, then the equation (4.2) simply is the equation (3.12).

In this thesis, the estimation of $\tilde{s}^*$ is done by the Bisection method. Recall in this method; the approximation of the actual root depends on the value of the tolerance ϵ which has been set for the algorithm. The algorithm runs when the values are greater than a defined tolerance until we achieve a given error ϵ, which corresponding to a theoretical error for s^*. Therefore after $\log_2\left(\frac{2}{100\epsilon}\right)$ iterations, we will have

$$|s^* - \widetilde{s}^*| \leq \epsilon, \quad |\xi(\widetilde{s}^*)| \leq \delta.$$

where δ is the absolute error for the function ξ.

Given that and using (4.2), we are expecting the following absolute error for the optimal value $\pi_{(2,\alpha)}$

with its estimated value $\widetilde{\pi}_{(2,\alpha)}$

$$\begin{aligned}\mathcal{E}_T := |\pi_{(2,\alpha)}(X) - \widetilde{\pi}_{(2,\alpha)}(X)| &\leq \frac{1}{1-\alpha}|\xi(s^*) - \xi(\widetilde{s}^*)| + (1+\frac{1}{1-\alpha})|s^* - \widetilde{s}^*| \\ &\leq \frac{1}{1-\alpha}\delta + (1+\frac{1}{1-\alpha})\epsilon. \end{aligned} \tag{4.3}$$

Note that here we neglecting the integral errors. In what follows, we set $\epsilon = 10^{-13}$ (for both the Bisection and integration methods) to calculate $\widetilde{s}^*$ and $\widetilde{\pi}_{(2,\alpha)}(.)$.

4.1.3 A Numerical Study

In this section, we carry out a numerical study base on our approximation technique to compute $\widetilde{\pi}_{(\alpha,2)}(.)$ for distributions that we discussed in Section 3.3.2. We compare the result with their exact values if the exact unique solution to the equation (3.9) is available. In this study, we considered two sides of the α levels; α and $1-\alpha$.

Case I: Uniform Distribution: $\mathcal{U}(0,1)$

In the special cases, $X \sim \mathcal{U}(0,1)$, one can find s^*, $\pi_{(p,\alpha)}$ as follows,

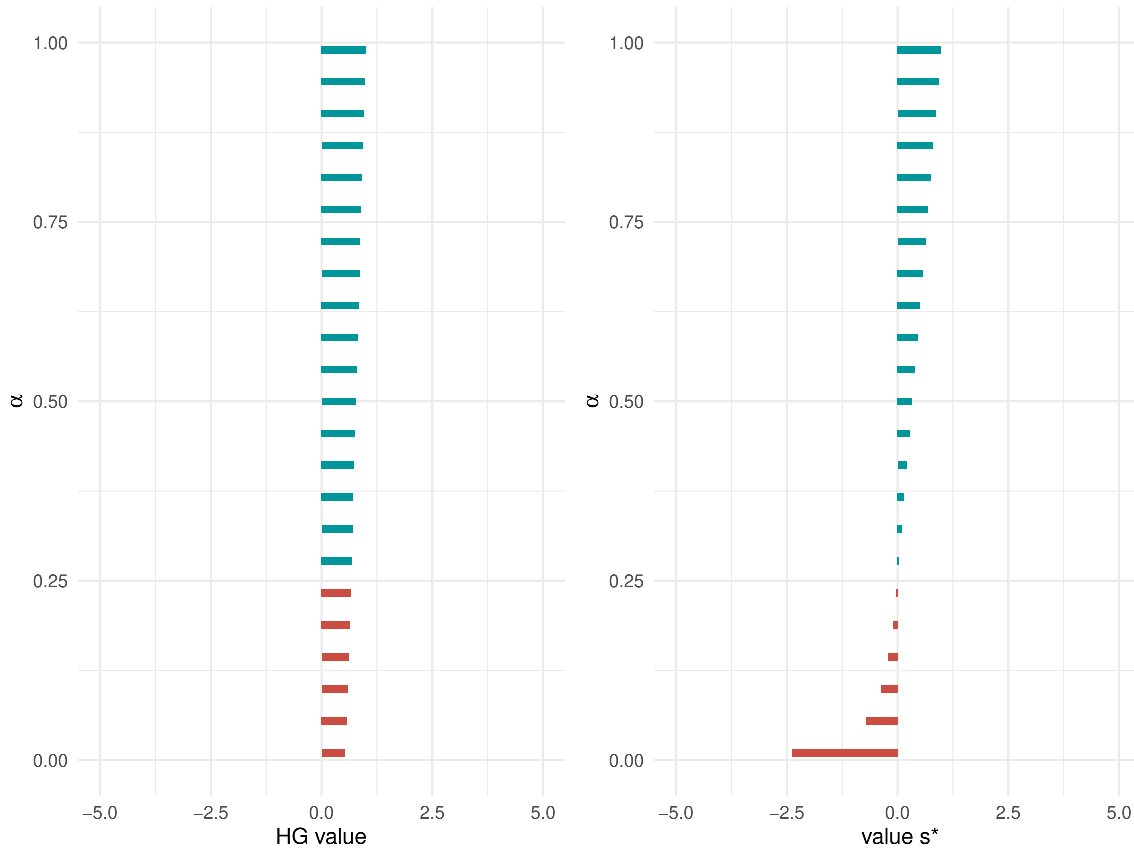
$$s^* = \begin{cases} \frac{1}{2}\left(1-\sqrt{\frac{1-\alpha}{3\alpha}}\right), & \alpha < \frac{1}{4}, \\ 0, & \alpha = \frac{1}{4}, \\ \frac{4}{3}\alpha - \frac{1}{3}, & \alpha > \frac{1}{4}, \end{cases} \quad \pi_{(2,\alpha)} = \begin{cases} \frac{1}{2}\left(1+\sqrt{\frac{\alpha}{3(1-\alpha)}}\right), & \alpha < \frac{1}{4}, \\ \frac{2}{3}, & \alpha = \frac{1}{4}, \\ \frac{4}{9}\alpha + \frac{5}{9}, & \alpha > \frac{1}{4}. \end{cases} \tag{4.4}$$

Table 4.1 reports the theoretical absolute error $\mathcal{E}_T$ (see, the equation (4.3)) and realized absolute error $\mathcal{E}_R$ at different levels of α when $X \sim \mathcal{U}(0,1)$. As this table demonstrates the numerical estimation of $\widetilde{\pi}_{(2,\alpha)}$ are very closed to the true values $\pi_{(2,\alpha)}$. One can also observe that the realized errors $\mathcal{E}_R$ are even better than the theoretical errors $\mathcal{E}_T$ in all cases of the α levels that have been studied.

α	$\widetilde{\pi}_{(2,\alpha)}$	$\pi_{(2,\alpha)}$	$\widetilde{s}^*$	s^*	$\mathcal{E}_T$	$\mathcal{E}_R$
When $\alpha < \frac{1}{4}$						
0.1	0.5962251	0.5962251	-0.3660254	-0.3660254	2.11357827E-13	0.00000000
0.050	0.5662266	0.5662266	-0.7583057	-0.7583057	2.05730620E-13	0.00000000
0.025	0.5462250	0.5462250	-1.3027756	-1.3027756	2.03475055E-13	1.11022302E-16
0.010	0.5290129	0.5290129	-2.3722813	-2.3722813	2.02804401E-13	1.11022302E-16
0.005	0.5204636	0.5204636	-3.5722639	-3.5722639	2.07643646E-13	3.33066907E-16
0.001	0.5091333	0.5091333	-8.6241438	-8.6241438	2.00100100E-13	0.00000000
When $\alpha > \frac{1}{4}$						
0.90	0.9555556	0.9555556	0.8666667	0.8666667	1.09999986E-12	0.00000000
0.95	0.9777778	0.9777778	0.9333333	0.9333333	2.09999986E-12	0.00000000
0.975	0.9888889	0.9888889	0.9666667	0.9666667	4.10000003E-12	0.00000000
0.990	0.9955556	0.9955556	0.9866667	0.9866667	1.01000000E-11	0.00000000
0.995	0.9977778	0.9977778	0.9933333	0.9933333	2.01000000E-11	0.00000000
0.999	0.9995556	0.9995556	0.9986667	0.9986667	1.00100000E-10	0.00000000

Table 4.1: This table illustrates the estimated value $\widetilde{\pi}_{(2,\alpha)}(X)$ and the actual value of $\pi_{(2,\alpha)}(X)$ where $X \sim \mathcal{U}(0,1)$. $\mathcal{E}_T$ and $\mathcal{E}_R$ are theoretical and realized absolute error respectively.

Figures 4.1b and 4.1a illustrate $\widetilde{s}^*$ and $\widetilde{\pi}_{(2,\alpha)}(.)$ (x-axis) vs values of α (y-axis). As expected by the equation (4.4), for any $\alpha < 0.25$, $\widetilde{s}^*$ takes negative values, and goes to zero as the α level goes to 0.25. For any $\alpha > 0.25$, s^* is positive. Note that the optimal value $\widetilde{\pi}_{(2,\alpha)}$ is always positive and increasing as a function of the α level.



(a) Estimated Value $\widetilde{\pi}_{(2,\alpha)}$ (b) Estimated Value $\widetilde{s}^*$

Figure 4.1: These figures show the optimal value $\tilde{\pi}_{(2,\alpha)}$ and the minimizer $\tilde{s}^*$ against α where $\alpha \in (0,1)$ when $X \sim \mathcal{U}(0,1)$.

Case II: Exponential Distribution: $Exp(1)$

By Example 48, when $X \sim Exp(1)$, we have

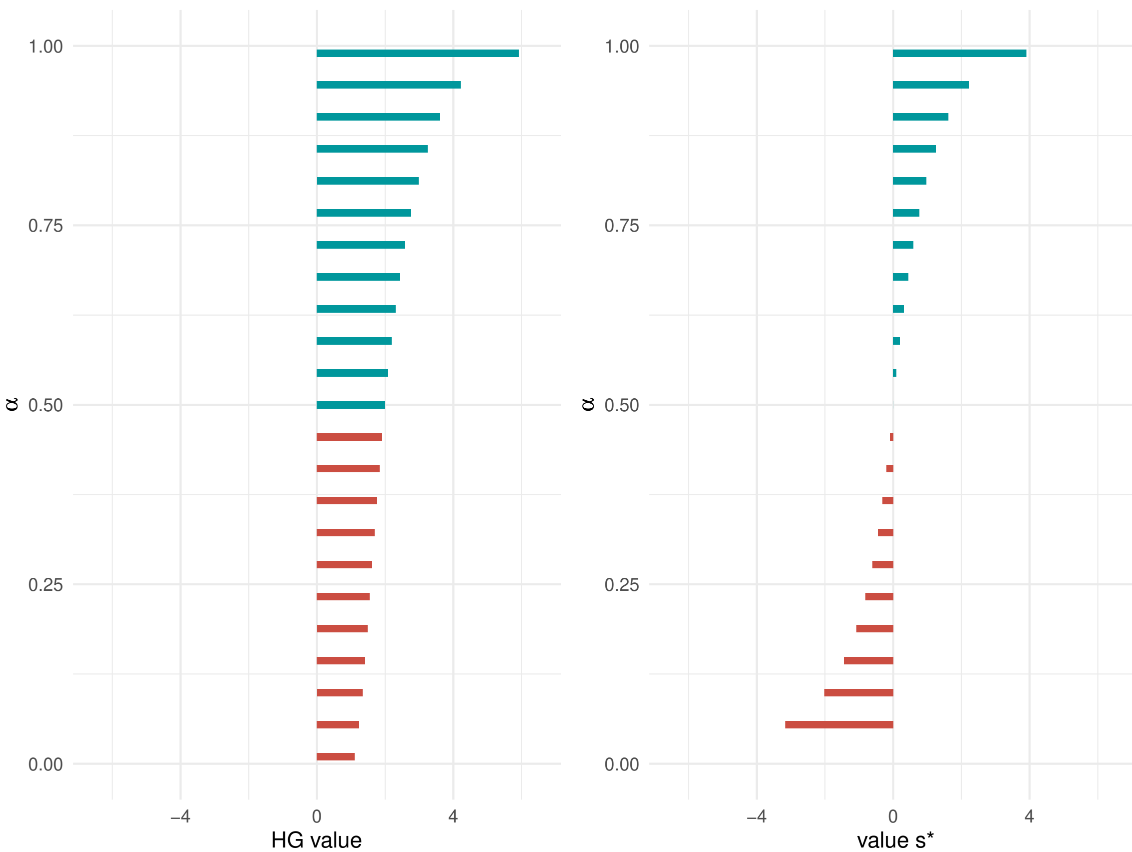
$$s^* = \begin{cases} -\ln(2-2\alpha), & \alpha > \frac{1}{2}, \\ 0 & \alpha = \frac{1}{2}, \\ 1 - \sqrt{\frac{1-\alpha}{\alpha}} & \alpha < \frac{1}{2}. \end{cases} \quad \pi_{(2,\alpha)} = \begin{cases} 2 - \ln(2-2\alpha) & \alpha > \frac{1}{2}, \\ 2 & \alpha = \frac{1}{2}, \\ 1 + \sqrt{\frac{\alpha}{1-\alpha}} & \alpha < \frac{1}{2}. \end{cases} \tag{4.5}$$

Table 4.2, lists the actual and approximation values of s^*, as well as the theoretical absolute error $\mathcal{E}_T$ and the realized absolute error $\mathcal{E}_R$ between $\pi_{(2,\alpha)}$ and $\widetilde{\pi}_{(2,\alpha)}$. Our observations from the table indicate that the numerical values for the risk measure $\pi_{(2,\alpha)}(.)$ are close to the true values, especially for any $\alpha \geq 0.5$. The numerical results show that, in most cases, the realized absolute error $\mathcal{E}_R$ is very insignificant. In general, the approximations are better when the α level is not closed to zero.

α	$\widetilde{\pi}_{(2,\alpha)}$	$\pi_{(2,\alpha)}$	$\widetilde{s}^*$	s^*	$\mathcal{E}_T$	$\mathcal{E}_R$
When $\alpha < \frac{1}{2}$						
0.10	1.3333333	1.3333333	-2.0000000	-2.0000000	2.17032301E-13	1.26565425E-14
0.05	1.2294157	1.2294157	-3.3588989	-3.3588989	2.16482254E-13	2.46469511E-14
0.025	1.1601282	1.1601282	-5.2449980	-5.2449980	2.02564103E-13	4.17443857E-14
0.010	1.1005038	1.1005038	-8.9498744	-8.9498744	1.86655702E-13	7.66053887E-14
0.005	1.0708881	1.0708881	-10.3106736	-10.3106736	3.14760641E-13	1.14797061E-13
0.001	1.0316386	1.0316386	-30.0606961	-30.0606961	2.00100100E-13	2.79110068E-13
When $\alpha > \frac{1}{2}$						
0.90	3.60943791	3.60943791	1.60943791	1.60943791	1.10000000E-12	0.00000000
0.95	4.30258509	4.30258509	2.30258509	2.30258509	2.10003469E-12	0.00000000
0.975	4.99573227	4.99573227	2.99573227	2.99573227	4.10006939E-12	0.00000000
0.99	5.91202301	5.91202301	3.91202301	3.91202301	1.01000271E-11	0.00000000
0.995	6.60517019	6.60517019	4.60517019	4.60517019	2.01000081E-11	8.88178420E-16
0.999	8.21460810	8.21460810	6.21460810	6.21460810	1.00099998E-10	1.77635684E-15

Table 4.2: This table shows the estimated value $\widetilde{\pi}_{(2,\alpha)}(X)$ and the actual value of $\pi_{(2,\alpha)}(X)$ where $X \sim Exp(1)$. $\mathcal{E}_T$ and $\mathcal{E}_R$ are theoretical and realized absolute error respectively.

Figures 4.2a and 4.2b plot $\widetilde{s}^*$ and $\widetilde{\pi}_{(2,\alpha)}$ (x-axis) against the α levels (y-axis). As point out in the equation (4.5), for any $\alpha < \frac{1}{2}$, $\widetilde{s}^*$ takes negative values, while is positive for any $\alpha > \frac{1}{2}$. As shown, at $\alpha = \frac{1}{2}$, $s^* = 0$ and $\pi_{(2,\alpha)} = 2$. The optimal value $\widetilde{\pi}_{(2,\alpha)}(.)$ is always positive and increasing as a function of α.



(a) Estimated Value $\widetilde{\pi}_{(2,\alpha)}$ (b) Estimated Value $\widetilde{s}^*$

Figure 4.2: These figures show the optimal value $\tilde{\pi}_{(2,\alpha)}$ and the minimizer $\tilde{s}^*$ against α where $\alpha \in (0,1)$.

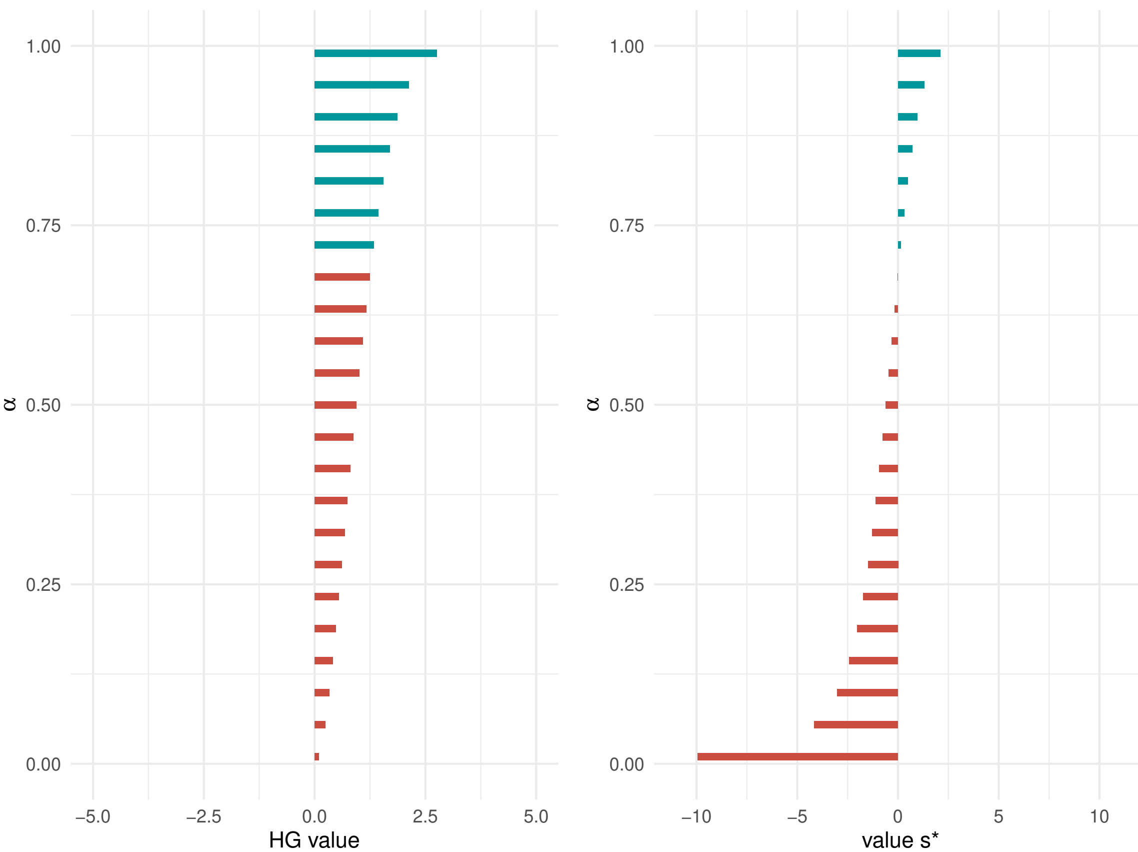
Case III: Normal Distribution: $\mathcal{N}(0,1)$

Finally, we investigate the estimation value of s^* when observations are contaminated by normally distributed function $X \sim \mathcal{N}(0,1)$. By proposition 50, we are able to solve the equation (3.41) numerically to find the values of $\widetilde{s}^*$ and $\widetilde{\pi}_{(2,\alpha)}$. Tables 4.3 presents the results of the estimation of $\widetilde{\pi}_{(2,\alpha)}$ at different levels of α with their theoretical absolute error $\mathcal{E}_T$. Note that, the upper bound for $\widetilde{\pi}_{(2,\alpha)}$ is calculated by using Corollary 51. However, we are not able to find the exact optimizer s^*, across all tested α, the results show that the approximation method can find the solution $\widetilde{s}^*$ and could be very close to the actual solution since, the absolute error $\mathcal{E}_T$ are very small in most cases.

α	$\widetilde{\pi}_{(2,\alpha)}$	upper bound for $\widetilde{\pi}_{(2,\alpha)}$	$\widehat{s}^*$	$\mathcal{E}_T$
When $\alpha < 1 - \frac{1}{\pi}$				
0.1	0.33329914	0.33336810	-2.99581381	2.13084841E-13
0.05	0.22941567	0.22941580	-4.35887029	2.12742555E-13
0.25	0.16012815	0.16012815	-6.24499800	2.09851720E-13
0.01	0.10050378	0.10050378	-9.94987437	2.15364500E-13
0.005	0.07088812	0.07088812	-10.41067360	2.00502513E-13
When $\alpha > 1 - \frac{1}{\pi}$				
0.9	1.86694006	2.28850983	0.948060801	1.10001735E-12
0.95	2.16604015	2.48243559	1.35249376	2.10000867E-12
0.975	2.43419317	2.67218003	1.69660145	4.10000651E-12
0.99	2.75423833	2.92555334	2.09081881	1.01000047E-11
0.995	2.97638568	3.11461725	2.35639349	2.01000020E-11
0.999	3.44309926	3.53458937	2.89836291	1.00100000E-10

Table 4.3: This table shows the estimated value $\widetilde{\pi}_{(2,\alpha)}(X)$ and the actual value of $\pi_{(2,\alpha)}(X)$ where $X \sim \mathcal{N}(0,1)$. $\mathcal{E}_T$ is the theoretical absolute error.

As one can observe from Figures 4.3, as expected, at any given level $\alpha < 1 - \frac{1}{\pi}$, the minimizer $\widetilde{s}^*$ is negative and $\widetilde{s}^* = 0$ at $\alpha = 1 - \frac{1}{\pi}$, with $\widetilde{\pi}_{(2,\alpha)}(X) = \sqrt{\frac{\pi}{2}}$. While for any $\alpha > 1 - \frac{1}{\pi}$, $\widetilde{s}^*$ is positive (see, Proposition 50). In general, $\tilde{s}^*$ is increasing as a function α.

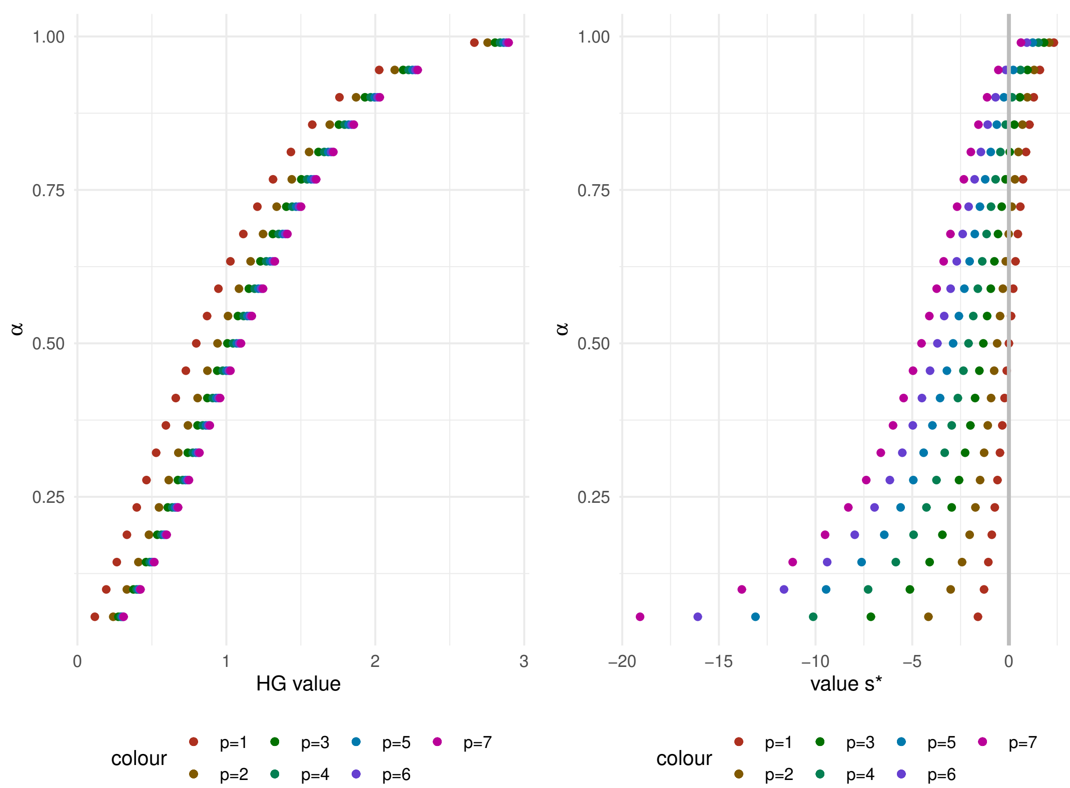


(a) Estimated Value $\widetilde{\pi}_{(2,\alpha)}$

(b) Estimated Value $\widetilde{s}^*$

Figure 4.3: These figures show the optimal value $\tilde{\pi}_{(2,\alpha)}$ and the minimizer $\tilde{s}^*$ when $\alpha \in (0,1)$, note that when $\alpha = 1 - \frac{1}{pi}$, $s^* = 0$.

For the final study of this section, we plot $\widetilde{\pi}_{(p,\alpha)}$ with different choices of p by considering $X \sim \mathcal{N}(0,1)$. The plot contained in Figure 4.4a demonstrates that $\widetilde{\pi}_{(p,\alpha)}$, with $p = 1,\ldots,7$ attain higher values as p increasing and $\widetilde{\pi}_{(p,\alpha)}$ always is the above of the $AV@R_\alpha$. Figure 4.4b shows that at a given $\alpha \in (0,1)$ as the value of p is increasing, the estimation values of $s^*_{p,\alpha}$ is increasing as well, while all minimizers $\widetilde{s^*}_{p,\alpha}$ are bounded by $\Phi^{-1}(\alpha)$. The gap between the values of $s^*_{p,\alpha}$ is reducing as the α level is goes to 1.



(a) Estimated Value $\widetilde{s}^*$

(b) Estimated Value $\widetilde{\pi}_{(2,\alpha)}$

Figure 4.4: Figures show the optimal value $\widetilde{s}^*$ (b) and $\widetilde{s}^*$ (a) as p increasing (labeled with different colors) vs α, note that when $\alpha = \frac{1}{2}$, then $\widetilde{s}^* = 0$, for any $\alpha > \frac{1}{4}$, $\widetilde{s}^* > 0$

Interestingly, as observed in Figure 4.4, with a fixed level of α, we graphically confirm, the minimizer $\tilde{s}^*_p$ is decreasing function of p, while $\widetilde{\pi}_{(p,\alpha)}$ is increasing as p increasing. Also as we show in Theorem 38, both $\widetilde{s}^*_p$ and $\widetilde{\pi}_{(p,\alpha)}$ are non-decreasing as a function α for any choice of p.

Finally, we should also point out; our program can handle to find the value of $s^*_{p,\alpha}$ numerically for many different kinds of distributions such as t-student and Pareto distribution and behave well on the power Young function with p rages from 1 to 7.

4.2 A Personal Case of Study: Empirical Study

A personal case study on the empirical application of the HG risk measure $\pi_{(p,\alpha)}$ on real data is done in this Section. To address comparison between the value of $\pi_{(p,\alpha)}(.)$ for different choice of p from a practical point of view, we analyze the Historical Simulation of three indices on the market. In the first subsection, we explain our methodology and our data-frame. In the second subsection, we employ the Historical Simulation method (Algorithm 2) as we discuss in the previous Chapter to study the performance of the $\pi_{(p,\alpha)}(.)$ in real data.

4.2.1 Data and Algorithm

In this subsection, we describe the methodology that we apply and the assumptions we make towards implementation of $\widehat{\pi}_{(p,\alpha)}(.)$ using the Historical Simulation algorithm.

The data series we investigate here include three stock market indices: Nasdaq Composite (US), Nikkei225 (Japan) and TSX (Canada) from January 1990 to December 2020, comprising 46728 observations. This daily historical data retrieves from Yahoo Finance. This time frame is chosen to view the economic phenomena' impact and also to have a large sample size for our investigation. Recall that most financial time-series studies focus on asset returns rather than asset prices. Here, we work with the (natural) logarithm of daily returns. Since we are interested in modelling the losses of the P&L, we multiply the returns by minus one. Therefore,

$$r_t = -\log\left(\frac{X_t}{X_{t-1}}\right) \times 100$$

where X_t is the daily closing value of stock indices on day t.

It should be noted that in the very few events when the price is missing, we employ an imputation procedure whereby the last recorded price of an earlier time is imputed, thus assuming the price is constant and the return is zero.

Algorithm 3: The Rolling Historical Simulation

For each data set, the value of $\widehat{\pi}_{(p,\alpha)}(.)$ is calculated base on the realized returns within the historical window. Note that the rolling window is the window length k or length of the data sub-sample.

Step 1: Let rolling window size be k and rolling window index i=1.

Step 2: Pick any $p \geq 1$ and a level of α, where $\alpha \in (0, 1)$.

Step 3: Compute $\widehat{\pi}^{(i)}_{(p,\alpha)}(.)$ (i.e. $\widehat{\pi}_{(p,\alpha)}(.)$ for i^{th}- window) for the $(k+i)^{th}$ day base on the observed log return $\{r_t\}_{i \leq t \leq k+i-1}$.

Step 4: To obtain $\widehat{\pi}_{(p,\alpha)}(.)$ for the next period, the window is moved one day forward. Let the rolling window index be i=i+1, and go to Step 5.

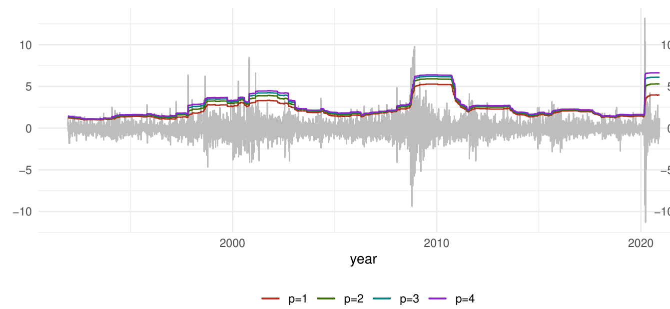
Step 5: Stop if $k+i-1$ is bigger than the length of data set N; otherwise, go to Step 3.

4.2.2 Results Analysis

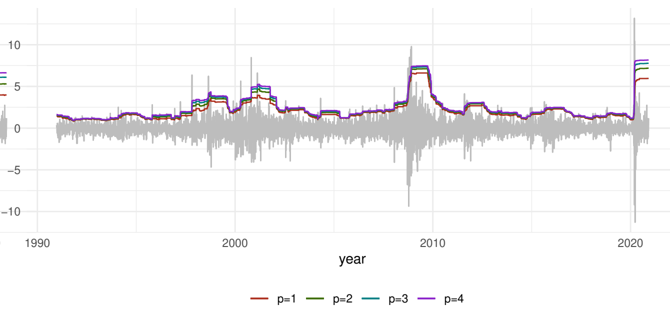
In this subsection, we provide an overview of estimating the $\pi_{(p,\alpha)}(.)$, using the rolling historical simulation algorithm on real data to explore the concepts of the HG risk measure in a more practical way. This empirical study made over thirty years. We consider four family members of the $\pi_{(p,\alpha)}(.)$, namely $AV@R_{\alpha}(.)$, $\pi_{(p,\alpha)}(.)$ with $p = 2, 3, 4$ and compare their historical simulation performance at the α levels 0.95 , 0.975, and 0.99. In addition, in this personal study, we try different window sizes, k, for the historical rolling window algorithm. First, we set the algorithm for window sizes of 502 days (two years), 251 days (one year), and 126 days (half a year) which are typical in the risk management field. Moreover, we apply a smaller sample size like 63 days (two months of observation) to examine the effect of the smaller sample size on the value of $\widehat{\pi}_{(p,\alpha)}(.)$.

4.2.2.1 A General Study

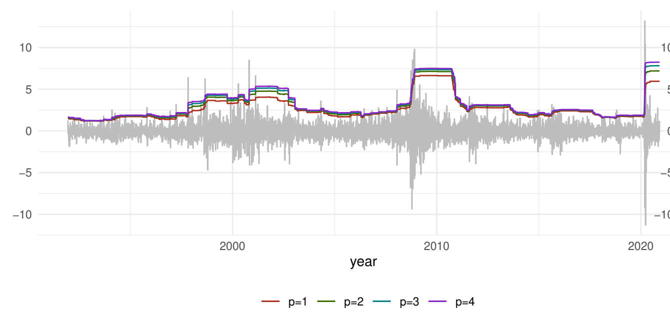
In the following, we expose plots that help to visualize $\widehat{\pi}_{(p,\alpha)}$, $p > 1$ in relation to the most common risk measure $\mathrm{AV@R}_{\alpha}$. As a first visualization, Figures 4.5-4.7, demonstrate the realized losses (negative logarithmic returns) for three indices, TSX, Nasdaq, Nikkei225 respectively, along with the estimated value of $\widehat{\pi}_{(p,\alpha)}$ with two sample size, two years and one year ($k = 502, 251$) at the α levels 0.95, 0.975, and 0.99. The plots contain in Figures 4.5, 4.6 and 4.7 demonstrate that $\widehat{\pi}_{(p,\alpha)}$ attains higher values when the value of p is increasing as which previously shown in Proposition 37 (iv). This behavior could be an evident of the benefit of using the $\pi_{(p,\alpha)},\ p > 1$ over use of the $\mathrm{AV@R}_\alpha$, especially in the stressed market, where it seems $\mathrm{AV@R}_\alpha$ is underestimated.



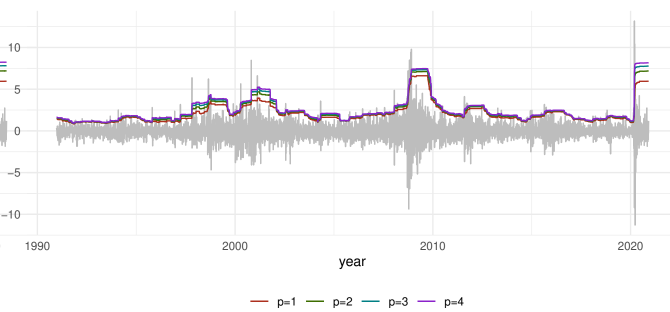
(a) $\widehat{\pi}_{(p,0.95)}$ with rolling window k=502



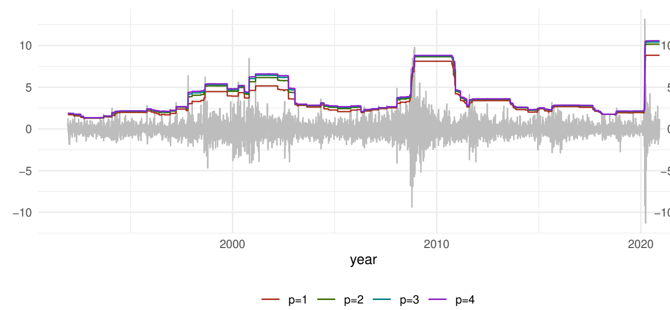
(b) $\widehat{\pi}_{(p,0.95)}$ with rolling window k=251



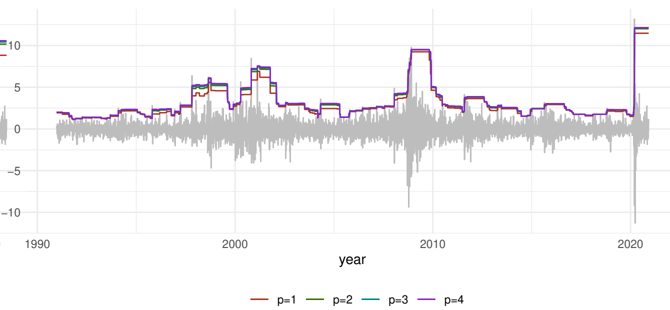
(c) $\widehat{\pi}_{(p,0.975)}$ with rolling window k=502



(d) $\widehat{\pi}_{(p,0.975)}$ with rolling window k=251

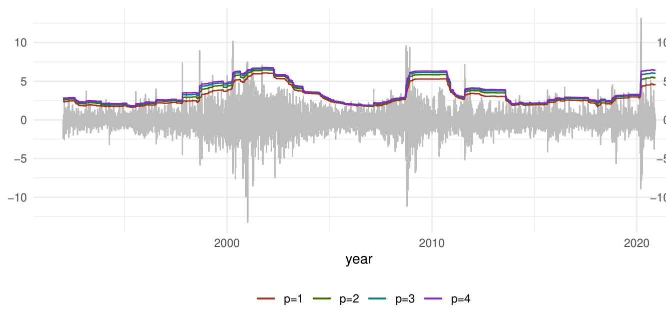


(e) $\widehat{\pi}_{(p,0.99)}$ with rolling window k=502

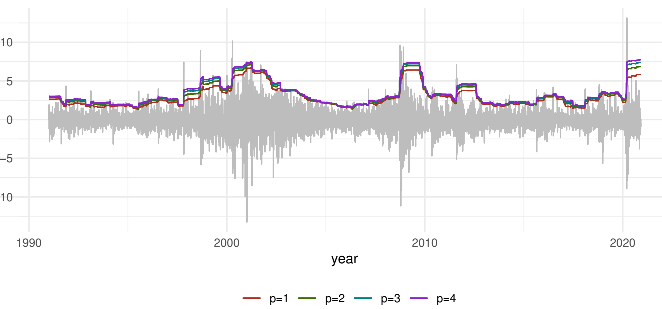


(f) $\widehat{\pi}_{(p,0.99)}$ with rolling window k=251

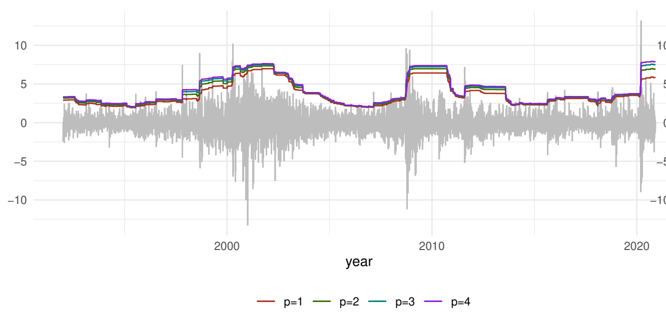
Figure 4.5: Daily log return of the TSX index from 1990 to 2020 (grey colour) and the estimated $\widehat{\pi}_{(p,\alpha)}$ with from $p = 1$ to $p = 4$, and $\alpha = 0.95,\ 0.975,\ 0.99$ over time with rolling window k=502 (the left side figures) and k=251 (the right side figures).



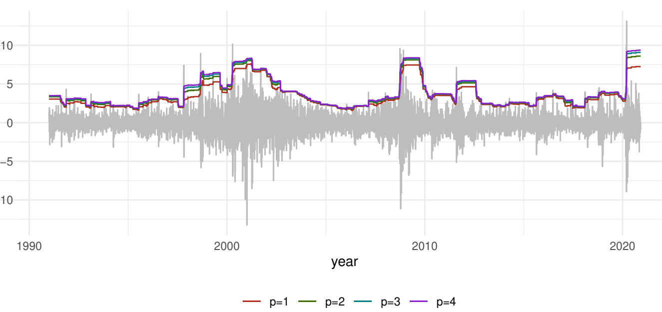
(a) $\widehat{\pi}_{(p,0.95)}$ with rolling window k=502



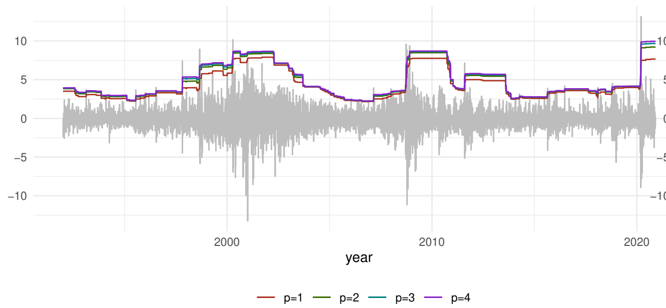
(b) $\widehat{\pi}_{(p,0.95)}$ with rolling window k=251



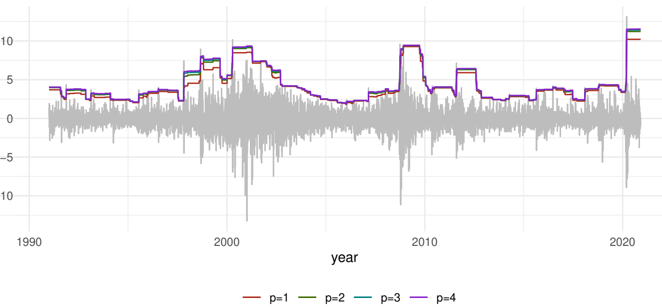
(c) $\widehat{\pi}_{(p,0.975)}$ with rolling window k=502



(d) $\widehat{\pi}_{(p,0.975)}$ with rolling window k=251

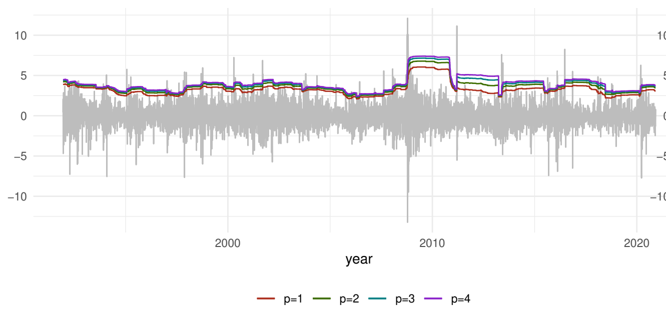


(e) $\widehat{\pi}_{(p,0.99)}$ with rolling window k=502

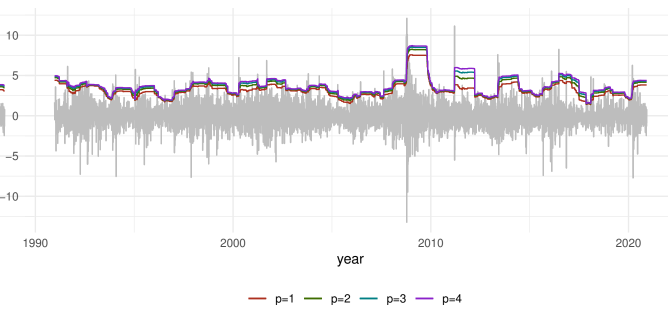


(f) $\widehat{\pi}_{(p,0.99)}$ with rolling window k=251

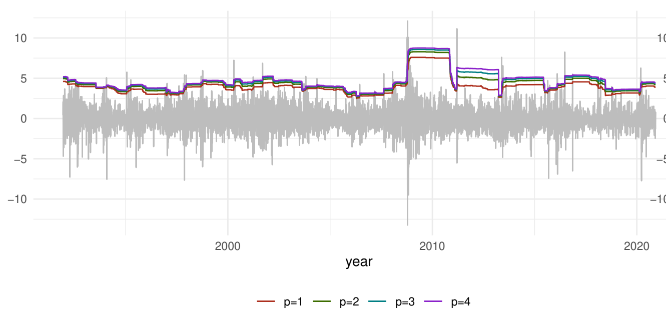
Figure 4.6: Daily log return of the Nasdaq index from 1990 to 2020 (grey colour) and the estimated $\widehat{\pi}_{(p,\alpha)}$ with from $p = 1$ to $p = 4$, and $\alpha = 0.95,\ 0.975,\ 0.99$ over time with rolling window k=502 (the left side figures) and k=251 (the right side figures).



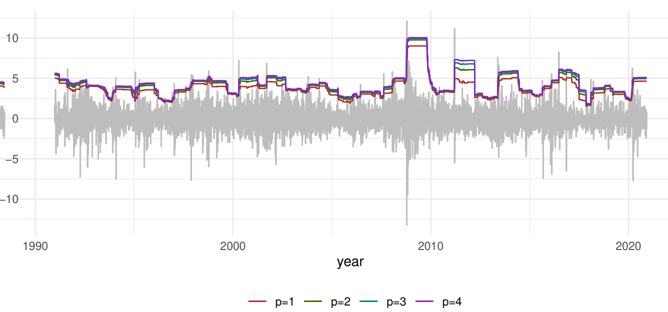
(a) $\widehat{\pi}_{(p,0.95)}$ with rolling window k=502



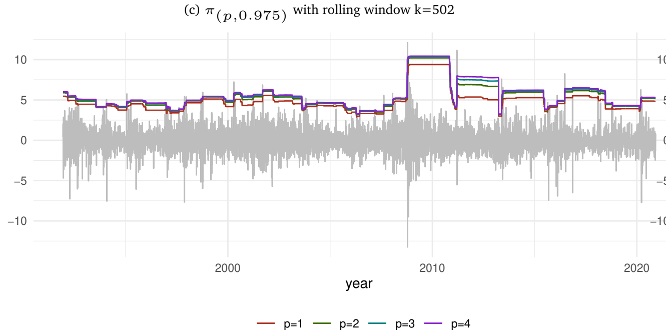
(b) $\widehat{\pi}_{(p,0.95)}$ with rolling window k=251



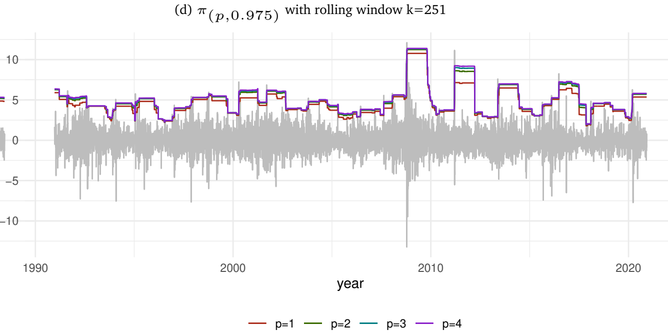
(c) $\widehat{\pi}_{(p,0.975)}$ with rolling window k=502



(d) $\widehat{\pi}_{(p,0.975)}$ with rolling window k=251



(e) $\widehat{\pi}_{(p,0.99)}$ with rolling window k=502

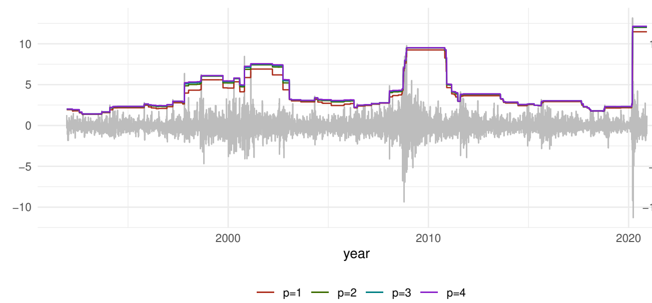


(f) $\widehat{\pi}_{(p,0.99)}$ with rolling window k=251

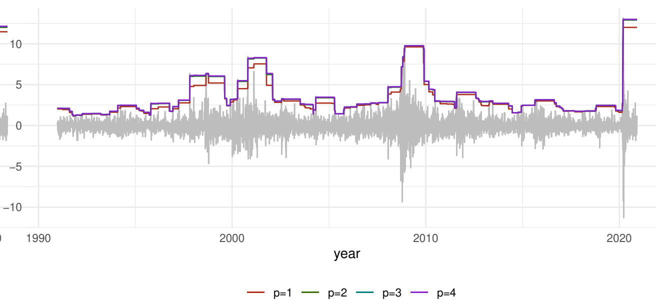
Figure 4.7: Daily log return of the Nikkei225 index from 1990 to 2020 (grey colour) and the estimated $\widehat{\pi}_{(p,\alpha)}$ with from $p = 1$ to $p = 4$, and $\alpha = 0.95,\ 0.975,\ 0.99$ over time with rolling window k=502 (the left side figures) and k=251 (the right side figures).

4.2.2.2 Effect of Higher Level of α

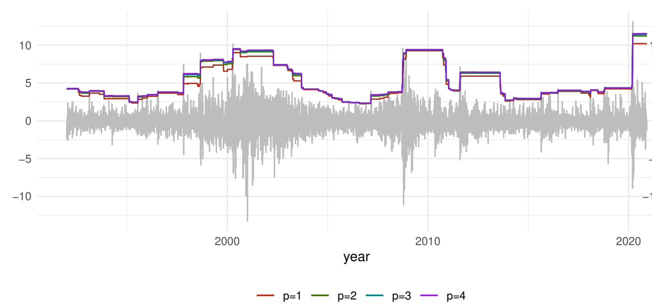
Note that, it is not unusual among financial institutions to measure risks at the higher α levels, e.g., 99.5% or 99.99%. In the following, we estimate $\widehat{\pi}_{(p,\alpha)}$ at the 99.5% level (see, figure 4.8). In this level, it is immediately clear than $\widehat{\pi}_{(p,0.995)},\ p > 1$ are more conservative than $\widehat{\mathrm{AV@R}}_{\alpha}$ (check Proposition 37 (iv)). Also one can observe that there is a clear difference between the value of the $\widehat{\mathrm{AV@R}}_{\alpha}$ and the other family members of $\widehat{\pi}_{(p,\alpha)}, p > 1$, especially in times of more volatile scenarios. Note that by Proposition 37 (vi), $\widehat{\pi}_{(p,\alpha)}$ is increasing function of α.



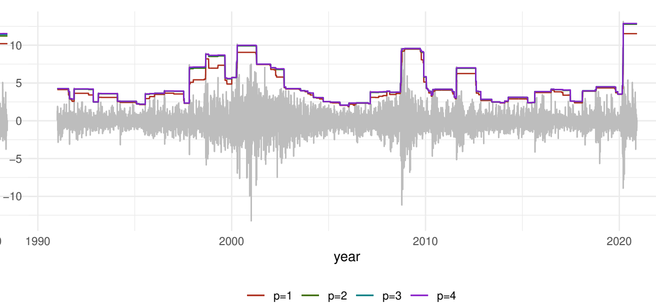
(a) TSX index, rolling window k=502



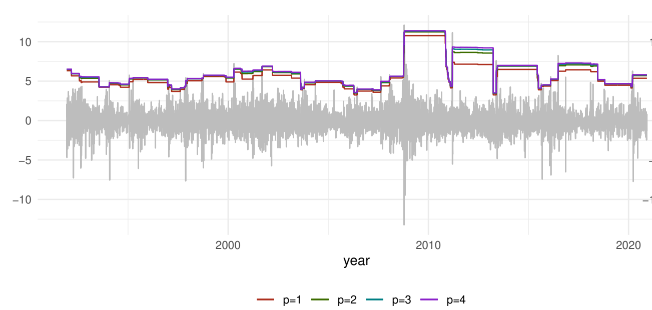
(b) TSX index, rolling window k=251



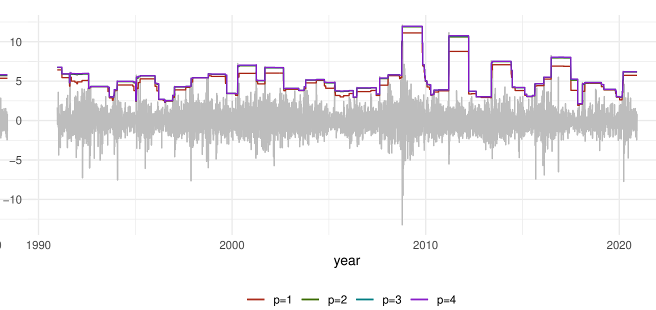
(c) Nasdaq index, rolling window k=502



(d) Nasdaq index, rolling window k=251



(e) Nikkei225 index, rolling window k=502



(f) Nikkei index, rolling window k=251

Figure 4.8: From top: TSX, Nasdaq and Nikkei indices from 1990 to 2020 and the estimated $\widehat{\pi}_{(p,995)}$ with $p = 1$ to $p = 4$, over time with rolling window k=502 (the left side figures) and k=251 (the right side figures).

4.2.2.3 Effect of Smaller Window Size

Practitioners often use window sizes between $k = 250$ and 1000. It seems interesting to see how this measure performs in shorter estimation window sizes. So, we update the window size with two new values, $k = 126, 63$ days and apply the rolling historical simulation algorithm for these sample sizes at the α level 0.95. As expected, in case of a short window size, the value of $\widehat{\pi}_{(p,\alpha)}$ is very sensitive to accidental outcomes from the recent past. A long window size, on the other hand, has the disadvantage that past data are included which might no longer be relevant to the current situation (see, figures 4.9 and 4.10).

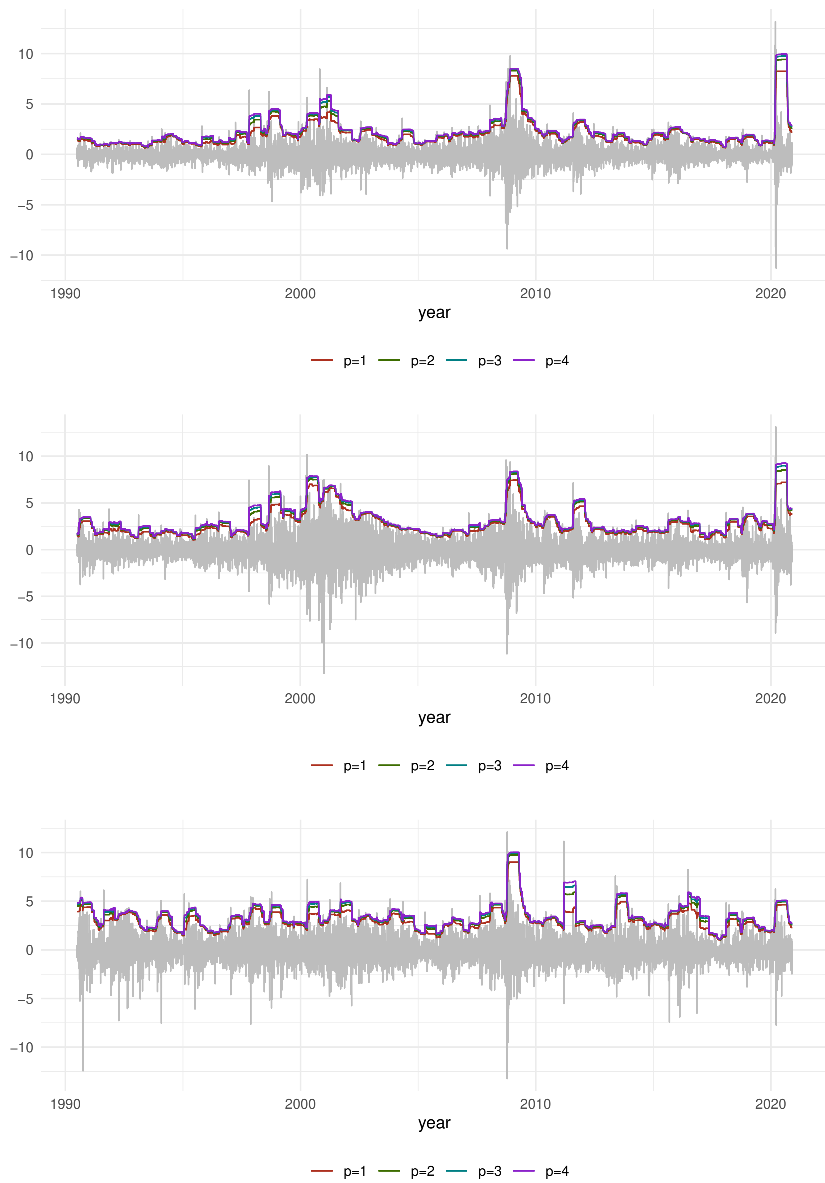


Figure 4.9: From top: TSX, Nasdaq and Nikkie returns and their estimated $\widehat{\pi}_{(p,0.95)}$ over time with rolling window k=126 days

In figure 4.10, one can observe that the value of $\widehat{\pi}_{(p,0.95)}(.)$ are enhanced significantly in the smaller window size, k=63, especially in extreme cases. Furthermore, it seems the HG risk measure with $p > 1$ performs better in a smaller sample size as it captures the tail events better.

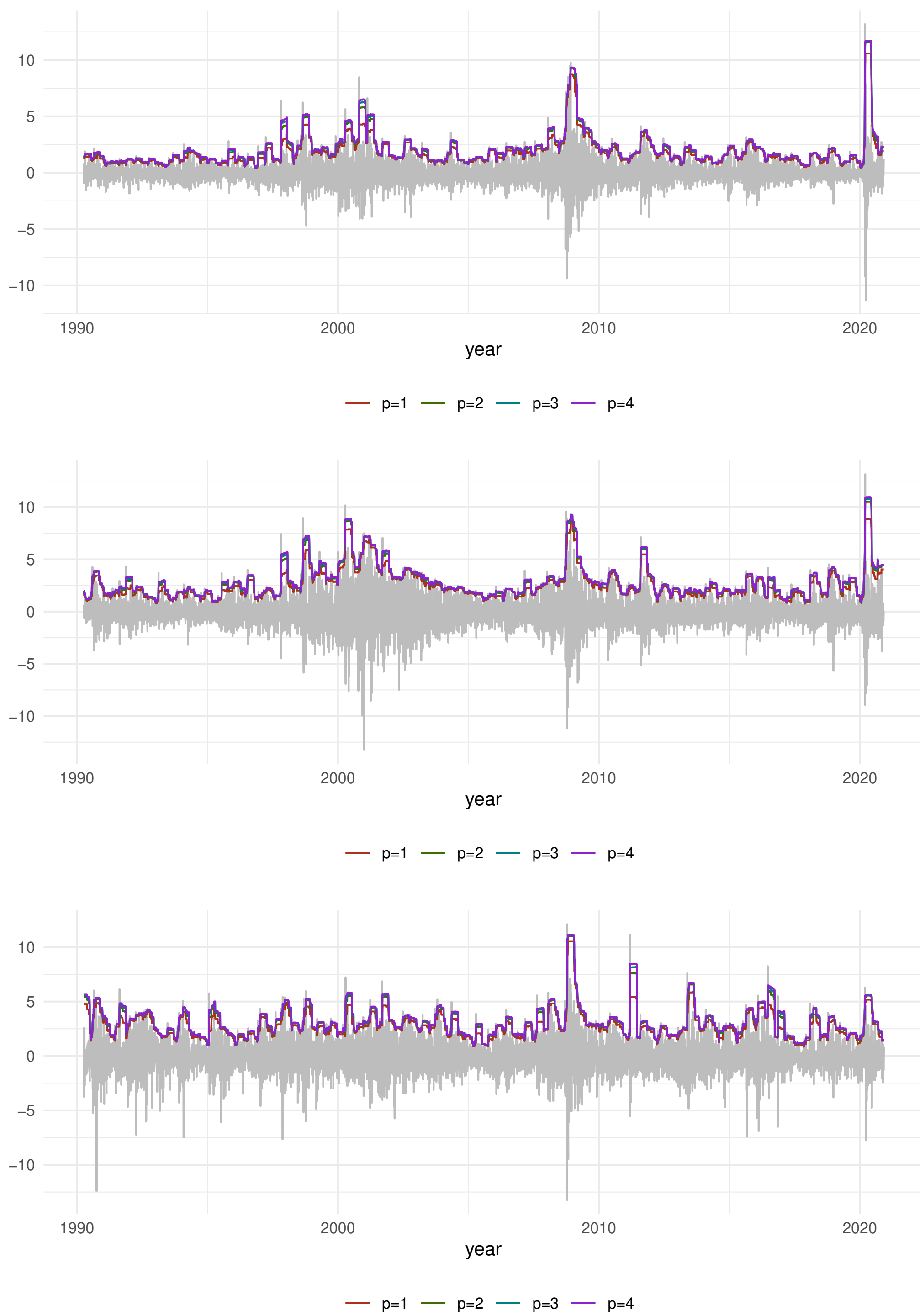


Figure 4.10: From top: TSX, Nasdaq and Nikkie returns and their estimated $\widehat{\pi}_{(p,0.95)}$ over time with rolling window k=63 days.

4.2.2.4 Effect of Standard Deviation and Kurtosis

Recall that the HG risk measure designs to put more weight on the tail and obviously fatter tail is associated with a higher risk in the fourth moment risk measure $\widehat{\pi}_{(4,\alpha)}$. Figure 4.11 displays the $\widehat{\mathrm{AV@R}}_{0.95}$ and the $\widehat{\pi}_{(4,.95)}$ as well as the historical standard deviation and the historical kurtosis with $k = 63$ days. The figure shows that periods of high kurtosis (fat tails) are reflected in measuring this risk; however, such an effect may not be present for the whole spectrum of time series. For instance, during the dot-com bubble, the kurtosis is lower than the measure (or those very close to it). The figure also suggests that the historical volatility or other factors may associate with the higher value of the HG risk measure along with the historical kurtosis.

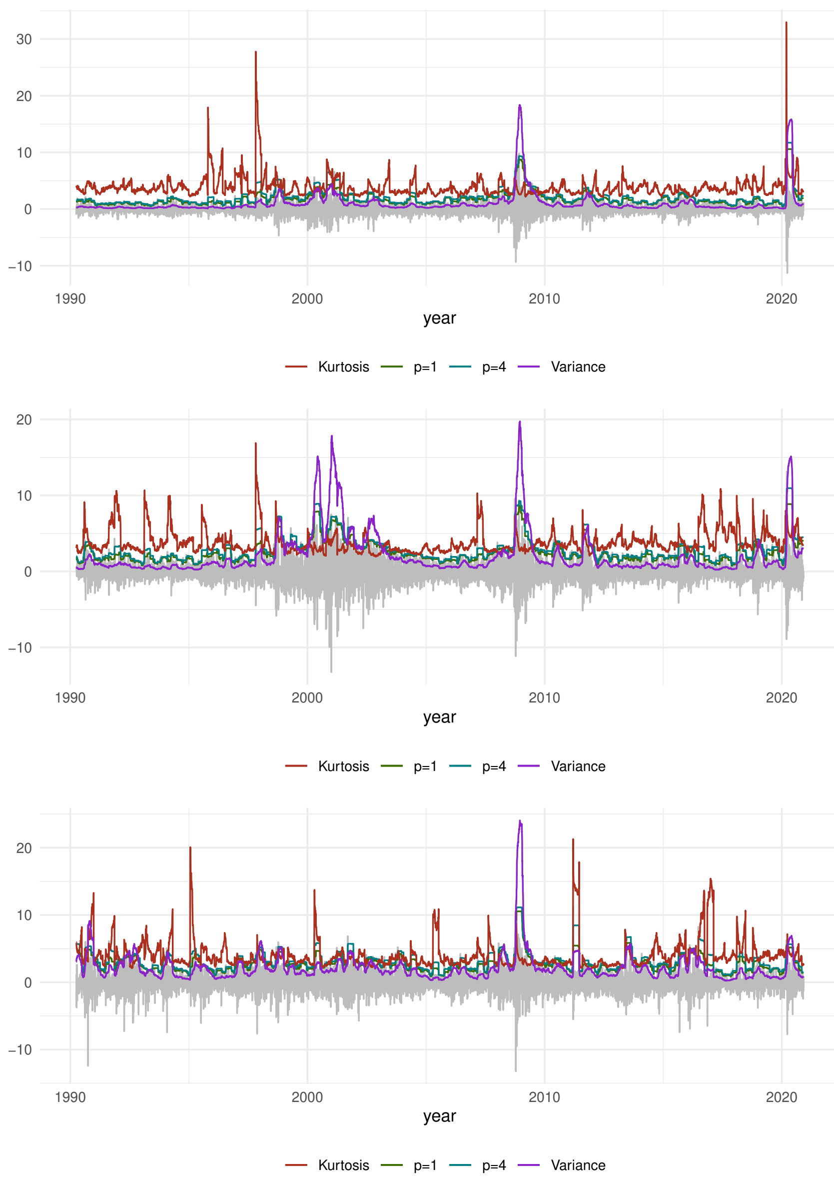


Figure 4.11: From top: TSX, Nasdaq and Nikkie returns and $\widehat{\pi}_{(1,.95)}$, $\widehat{\pi}_{(4,.95)}$ along the historical kurtosis, volatility with rolling window k=63 days.

4.2.3 Summary

To summarize our findings, we could say a few broad observations are obtained by this study's outcomes. This historical simulation study shows, especially in some financial market turbulence such as the dot-com bubble (1995-2002), the global financial crisis (2007-2009) and the COVID-19 recession (2020), during which the $\mathrm{AV@R}_{\alpha}$ seems (severely) underestimated, the empirical estimation of the $\pi_{(p,\alpha)},\ p>1$ is more conservative than the $\mathrm{AV@R}_{\alpha}$, even for shorter estimation windows such as 63 days (see, Figure 4.10).

As we expected, the higher the confident α levels, the sensitivity of $\widehat{\pi}_{(p,\alpha)},\ p\geq 1$ is higher because of capturing more extreme situations. In general, ignoring some extreme circumstances, at the higher α levels such as 99%, and 99.5%, it is difficult to say if there are any clear differences between the different choice of $\pi_{(p,\alpha)}$ for $p>1$ (Figure 4.8). Interestingly enough, the value of the $\widehat{\pi}_{(p,\alpha)},\ p>1$ at these levels are fairly closed and it seems apparently there is no different between $s^{*}_{p,\alpha}$ and $s^{*}_{p-1,\alpha}$ at the higher α levels as p increasing. While when $p=1$, one can observe that the $\widehat{\mathrm{AV@R}}_{\alpha}$ is significantly smaller than the other family members $\widehat{\pi}_{(p,\alpha)}$ with $p>1$. Therefore, we can recommend that it may be beneficial to use the $\widehat{\pi}_{(2,\alpha)}$ over $\widehat{\mathrm{AV@R}}_{\alpha}$ which is not computationally expensive as the other choice of p with fixed α when more conservative risk measure is needed

At a given α level, depending on the choice for the window size, the value of $\widehat{\pi}_{(p,\alpha)}$ may be volatile: large windows keep data points for a long time which can lead to sustained periods of constant $\widehat{\pi}_{(p,\alpha)}$ for any $p\geq 1$, while small windows eject data points relatively fast resulting in swift changes in the value of $\widehat{\pi}_{(p,\alpha)}$. According to the observation, even though there is no evidence, as the sample size is decreasing there is not significantly different between the historical estimation of $\pi_{(p,\alpha)}$ while p varies from 2 to 4. On the other hand, for large sample sizes, it seems $\widehat{\pi}_{(p,\alpha)}$ with larger p, is more conservative (Figures 4.5, 4.6, 4.7).

Finally, to re-emphasize, the personal study in this chapter is further supported by the fact that the HG risk measure ($\pi_{(p,\alpha)}\ p=2,3,4$) is similar in the structure and characteristics to the $\mathrm{AV@R}_{\alpha}$ ($\pi_{(p,\alpha)},\ p=1$). However, the obtained results of this section are data-driven; the presented preliminary case studies indicate that the developed family $\pi_{(p,\alpha)}$, $p>1$ demonstrate auspicious performance and could be a better indicator of risk measures, especially in extreme circumstances where it seems the $\mathrm{AV@R}_{\alpha}$ is underestimated.

Appendices

Appendix A

Preliminary

For convenience of reader, in this appendix, we recall a few basic concepts on topological spaces and the mathematical foundations of probability theory which be used throughout this thesis. For further facts, we refer the reader to [16, 60].

A.1 Topology: Basic Definitions and Theorems

Definition 59. *(Topology) Given a set X a topology τ on X is a collection of subsets of X, such that :*

(i) $\emptyset, X \in \tau$.

(ii) Arbitrary unions of sets in τ are in τ.

(iii) Finite intersections of sets in τ are in τ.

A set X with a topology τ is called a topological space, and is denoted (X, τ), or simply X when no confusion. We call a member of τ an open set in X.

A topological space X is metrizable if there exists a metric d on X that generates the topology of X.

Definition 60. *(Relative topology) Let (X, τ) be a topological space and let $E \subset X$. Then $\tau \cap E$ is called the relative topology or the topology induced by τ on E, and collection τ_E of subsets of E, defined by*

$$\tau_E := \{V \cap E : V \in \tau\},$$

is a topology on E.

When $E \subset X$ is equipped with its relative topology, we call E a (topological) subspace of X. A set in τ_E is called (relatively) open in E.

Nets

A net is like a sequence, except that instead of being indexed by the natural numbers, the index set can be more general. In particular, sequences are nets.

Definition 61. *A net in a set X is a function $x : D \to X$, where D is a directed set. The directed set D is called the index set of the net and the members of D are indexes.*

We denote the function $x(.)$ by $(x_\alpha)_{\alpha \in D}$ and we write $(x_\alpha)_{\alpha \in D} \subset X$ for net $(x_\alpha)_{\alpha \in D}$ in X.

In a topological space (X, τ), and a given point x of X, we call a subset N of X a neighbourhood of X if we can find an open set O such that $x \in O \subseteq N$.

Definition 62. *A net (x_α) in a topological space (X, τ) converges to some point x if for each neighborhood V of x there exists some index α_0 (depending on V) such that $x_\alpha \in V$ for all $\alpha \geq \alpha_0$. We say that x is the limit of the net, and write*

$$x_\alpha \to x \quad \text{or} \quad x_\alpha \xrightarrow{\tau} x.$$

Note that in a metric space, $x_\alpha \to x$ if and only if $d(x_\alpha, x) \to 0$.

Definition 63. *(Subnet) A net $(y_\beta)_{\beta \in I}$ is a subnet of a net $(x_\alpha)_{\alpha \in J}$ if there is a function $f : I \to J$ satisfying*

(i) $x_\beta = x_{f_\beta}$ for each $\beta \in I$, where $f_\beta := f(\beta)$.

(ii) For each $\alpha_0 \in J$, they exists some $\beta_0 \in I$ such that $\beta \geq \beta_0$ implies $f_\beta \geq \alpha_0$.

In the topological space, a net converges to a point if and only if every subnet converges to that same point (see, for example, [16], Lemma 2.17).

Continuous Functions

Definition 64. *A function $\varphi : X \to Y$ between topological space is continuous if $\varphi^{-1}(U)$ is open in X for each open set U in Y.*

Note that a function $\varphi : X \to Y$ is continuous at x, if and only if a net $x_\alpha \to x$ in X, then $\varphi(x_\alpha) \to \varphi(x)$ in Y ([16], Theorem 2.28).

The notion of lower semicontinuity can be extended in a straightforward manner to topological spaces.

Definition 65. *Let τ be a topology on X. We say that $\varphi : X \to [-\infty, \infty]$ is τ* lower semicontinuous*, whenever the sublevel set*

$$\{\varphi \leq \lambda\} := \{x \in X : \varphi(x) \leq \lambda\}$$

is τ closed for each $\lambda \in \mathbb{R}$.

The next lemma (see, [16] for more detail) is a characterization of lower semicontinuity which sometimes used as a definition.

Lemma 66. *([16], Lemma 2.42)) Let $\varphi : X \to [-\infty, \infty]$ be a function on a topological space. Then φ is lower semicontinuous if and only if for any net $(x_\alpha) \subset X$,*

$$x_\alpha \xrightarrow{\tau} x$$

implies

$$\varphi(x) \leq \liminf_{\alpha} \varphi(x_\alpha).$$

Note that that if τ is a metrizable topology, nets can be replaced by sequences, so the above is holding for any sequence.

Weak Topology

Given a topology τ, we can determine whether a function is continuous. This argument can be reversed: given a space and functions on that space, we can define a topology such that those functions are continuous. Let X be a nonempty set, and $\{(Y_i, \tau_i)\}_i$ be a family of topological spaces and for each i let $\varphi_i : X \to Y_i$ be a function. The weak topology or initial topology on X generated by the family of functions $(\varphi_i)_i$ is the weakest topology on X that makes all the functions φ_i continuous. It is the topology generated by the family of sets

$$\{\varphi_i^{-1}(V) : \ V \in \tau_i\}.$$

An important special case is a weak topology generated by a family of real functions. For a family X^* of real functions on X, the weak topology generated by X^* is denoted $\sigma(X, X^*)$.

Let X^* be a family of real-valued functions on a set X. As we indicated before, every subset $A \subset X$ has a relative topology induced by the $\sigma(X, X^*)$ weak topology on X. It also has its own weak topology, the $\sigma(A, X^*|_A)$ topology, where $X^*|_A$ is the family of restrictions of the functions in X^* to A.

Lemma 67. *(Relative weak topology, [16], 2.53) Let X^* be a family of real-valued functions on a set X, and let A be a subset of X. The $\sigma(A, X^*|_A)$ weak topology on A is the relative topology on A induced by the $\sigma(X, X^*)$ weak topology on X.*

Definition 68. *(Weakly convergent) $(x_\alpha)_\alpha$ in X converges weakly to $X \in X$, denoted by*

$$x_\alpha \xrightarrow{\sigma(X,\mathcal{X}^*)} x$$

if and only if

$$\varphi(x_\alpha) \to \varphi(x), \quad \textit{for each} \ \ \varphi \in X^*.$$

We conclude this section by recalling the conjugate of $\varphi : X \to [-\infty, \infty]$. To this effect, recall that the conjugate of φ is the functional $\varphi^* : X^* \to [-\infty, \infty]$ defined by

$$\varphi^*(y) := sup_{x \in X}\{< y, x > -\varphi(x)\}, \quad \text{for all } y \in X^*.$$

where the symbol $<,>$ indicates the bilinear form for dual (X, X^*).

A.2 Measures: Basic Definitions and Theorems

Measurable Spaces

Intuitively, the sample space Ω is a set of all possible outcomes $\omega \in \Omega$ of some random experiment. The event space $\mathcal{F}$ represents both the amount of information available as a result of the experiment conducted and the collection of all subsets of possible interest to us, where we denote elements of $\mathcal{F}$ as events. In finance, it is important to know which information is available to investors. This is formalized using σ-algebras.

Definition 69. *(σ-algebra) A collection $\mathcal{F}$ of subsets of Ω is called a σ-algebra (or σ-field) if the following hold.*

(i) $\emptyset \in \mathcal{F}$.

(ii) If $A \in \mathcal{F}$ *then* $A^c \in \mathcal{F}$.

(iii) $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$ *whenever* $A_n \in \mathcal{F}$, *for all* $n \in \mathbb{N}$.

A pair $(\Omega, \mathcal{F})$ with $\mathcal{F}$ a σ-algebra of subsets of Ω is called a measurable space.

Definition 70. *(A measure) Given a measurable space* $(\Omega, \mathcal{F})$. *A measure (finitely additive measure) is a function* $\mu : \mathcal{F} \to [0, \infty]$ *such that*

(i) $\mu(\emptyset) = 0$.

(ii) $\mu(\sum_n A_n) = \sum_n \mu(A_n)$ *for countable (finite) disjoint sequences* A_n *in* $\mathcal{F}$.

Remark 71.

(i) μ *is finite measure if* $\mu(\Omega) < \infty$, *and is an infinite measure if* $\mu(\Omega) = \infty$.

(ii) μ *is a probability measure if* $\mu(\Omega) = 1$, *and often labeled by* $\mathbb{P}$.

(iii) The sets $N \in \mathcal{F}$ *such that* $\mu(N) = 0$ *are called null sets.*

(iv) A property holds almost surely (a.s.) or for almost all $w \in \Omega$ *everywhere on* $\Omega \setminus N$, *where* N *is a null set.*

A measure space is a triplet $(\Omega, \mathcal{F}, \mu)$, *with* μ *a measure on the measurable space* $(\Omega, \mathcal{F})$. *A probability space is measure space* $(\Omega, \mathcal{F}, \mathbb{P})$ *with* $\mathbb{P}$ *a probability measure.*

Definition 72. *A probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *is non-atomic, or alternatively call* $\mathbb{P}$ *non-atomic if for any* $A \in \mathcal{F}$, *and* $\mathbb{P}(A) > 0$, *then there exists* $B \in \mathcal{F}$ *such that* $B \subset A$ *with* $0 < \mathbb{P}(B) < \mathbb{P}(A)$.

Note that if a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is non-atomic and $A \in \mathcal{F}$ with $\mathbb{P}(A) > 0$, then If $0 < a < \mathbb{P}(A)$ then there exists $B \subset A$ with $\mathbb{P}(B) = a$.

Definition 73. *Let* Ω *be a non-empty set. Given a collection of subsets* $A_\alpha \subseteq \Omega$ *(not necessarily countable). we denote the smallest* σ*-algebra* $\mathcal{F}$ *such that* $A_\alpha \in \mathcal{F}$ *for all* $\alpha \in \Gamma$ *either by* $\sigma(\{A_\alpha\})$ *or by* $\sigma(A_\alpha, \alpha \in \Gamma)$, *and call* $\sigma(A_\alpha, \alpha \in \Gamma)$ *the* σ*-algebra generated by the sets* A_α.

$$\sigma(A_\alpha, \alpha \in \Gamma) := \bigcap \{\mathcal{G} : \mathcal{G} \text{ is a } \sigma\text{-algebra}, A_\alpha \in \mathcal{G}, \text{ for all } \alpha \in \Gamma\}.$$

The most important example of σ-algebra is the σ-algebra of subsets of a topological space generate by its open sets.

Definition 74. *The Borel* σ*-algebra of a topological space* (X, τ), *denoted by* $\mathcal{B}$ *is* $\sigma(\tau)$, *the* σ*-algebra generated by family* τ *of open sets. Members of Borel* σ*-algebra are Borel sets.*

Unless otherwise stated, The real line $\mathbb{R}$ is equipped with the Borel σ-algebra.

Measurable Functions

Definition 75. *If $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ are two measurable spaces, a function $X : \Omega_1 \to \Omega_2$ is called measurable if for any set $B \in \mathcal{F}_2$, the set*

$$[X \in B] = X^{-1}(B) := \{\omega \in \Omega : X(\omega) \in B\},$$

is in $\mathcal{F}_1$.

Given a function $X : \Omega \to \mathbb{R}$, we denote by $\sigma(X)$ the smallest σ-algebra $\mathcal{F}$ such that X is a measurable mapping from $(\Omega, \mathcal{F})$ to $(\mathbb{R}, \mathcal{B})$. Alternatively,

$$\sigma(X) := \sigma(\{\omega : X(\omega) \le \alpha\}, \alpha \in \mathbb{R}) = X^{-1}(\mathcal{B}).$$

For example, if $X = \mathbb{1}_A$, then $X^{-1}(\{-1\}) = A$, $X^{-1}(\{0\}) = A^c$ and $\sigma(X) = \{\emptyset, \Omega, A, A^c\}$.

Definition 76. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, a real-valued function $X : \Omega \to \mathbb{R}$ is called a random variable if it is a measurable function when considered as a function from the measurable space $(\Omega, \mathcal{F})$ to the measurable space $(\mathbb{R}, \mathcal{B})$, where $\mathcal{B}$ is the Borel σ-algebra on $\mathbb{R}$. Namely,*

$$X^{-1}((x, \infty)) := [X > x] \in \mathcal{F}, \quad \textit{for all } \; x \in \mathbb{R}.$$

Distribution function

As defined next, every random variable X induces a probability measure on its range which is called the law of X.

Definition 77. *The law (distribution) of a real-valued random variable X, denoted law$_X$[1], is the probability measure on $(\mathbb{R}, \mathcal{B})$ such that*

$$\textit{law}_X(B) := \mathbb{P}(X \in B) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in B\}), \quad \textit{for all } B \in \mathcal{B}.$$

The number $\mathbb{P}(X \in B)$ is the probability that X “falls in B” (or “takes its value in B”).

Definition 78. *X equals Y in law (or in distribution), denoted by $X \sim Y$, if and only if*

$$\textit{law}_X(B) = \textit{law}_Y(B), \quad \textit{for any Borel set } B.$$

Instead of working with distributions of random variables, which are probability measure on the measurable space $(\mathbb{R}, \mathcal{B})$, one can encode them in a simpler object called a distribution function (sometimes referred to as a cumulative distribution function, or c.d.f.)

[1] Sometime is denoted by $\mathbb{P} \circ X^{-1}(B) := \mathbb{P}[X \in B]$

Definition 79. *The cumulative distribution function of a random variable* X *defined on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *is the function* $F_X : \mathbb{R} \to [0,1]$ *defined by*

$$F_X(x) := \mathbb{P}(X \leq x) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \leq x\}), \quad x \in \mathbb{R}.$$

and its survival function by

$$\bar{F}_X(x) := \mathbb{P}(X > x).$$

Definition 80. *A function* $F : \mathbb{R} \to [0,1]$ *satisfying*

(i) F *is right continuous,*

(ii) F *is monotone non-decreasing,*

(iii) $\lim_{x\to\infty} F(x) = 1$*,* $\lim_{x\to-\infty} F(x) = 0$*,*

is call a (probability) distribution function.

Given a function F, the corresponding quantile function F^{-1} is a real-valued left continuous function on $(0, 1)$ defined as

$$F^{-1}(\alpha) := \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}.$$

The quantile function can be seen as some kind of an inverse of the cdf of F. Indeed, F^{-1} is the ordinary inverse if F is strictly increasing. When the cdf F is continuous and strictly increasing, the quantile function F^{-1} is exactly equal to inverse of F.

A.3 L^p Spaces: Basic Definitions and Theorems

Throughout this thesis, we assume that the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is non-atomic (i.e. is rich enough to support a random variable with a continuous distribution). Ω is a collection of the possible states of the world, $\mathcal{F}$ is a σ-algebra on Ω and $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$. All equalities and inequalities between random variables are understood in the $\mathbb{P}$-almost sure (a.s.) sense.

Expectation

One of the most fundamental concepts of probability theory is the expectation of a random variable. Many integration results are proven by first showing they hold for simple functions and then extending the result to a more general function.

Definition 81. *(Characteristic function) Let* $(\Omega, \mathcal{F})$ *be a measurable space. The characteristic function of a set* $A \in \mathcal{F}$ *is defined as being*

$$\mathbb{1}_A : \Omega \to \mathbb{R},$$

where

$$\mathbb{1}_A(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & o.w. \end{cases}$$

Definition 82. *(Simple function) A simple function is a function $X : \Omega \to \mathbb{R}$ of the form*

$$X(\omega) = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}(\omega).$$

such that $a_i \in \mathbb{R}$, $A_i \in \mathcal{F}$ is a sequence of disjoint and $\bigcup_{i=1}^{n} A_i = \Omega$.

Let $\mathcal{L}$ be the set of all simple functions on Ω. The set $\mathcal{L}$ has the following properties of $\mathcal{L}$

(i) $\mathcal{L}$ is a vector space. This means if

$$X = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i} \in \mathcal{L}, \quad \text{and} \quad Y = \sum_{j=1}^{m} b_j \mathbb{1}_{B_j} \in \mathcal{L},$$

then

$$\alpha X + Y = \sum_{i,j} (\alpha a_i + b_j) \mathbb{1}_{A_i B_j}.$$

(ii) if $X, Y \in \mathcal{L}$, then $XY \in \mathcal{L}$ since

$$XY = \sum_{i,j} a_i b_j \mathbb{1}_{A_i \cap B_j}.$$

Note that, any positive random variable can be approximated by a sequence of simple function. Namely, any positive random variable X is $\mathcal{F}$-measurable,if and only if there is an increasing sequence $(X_n)_{n\in\mathbb{N}}$ of simple function where

$$X_n := \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \mathbb{1}_{\left\{\frac{k-1}{2^n} \le X \le \frac{k}{2^n}\right\}} + n\mathbb{1}_{\{X>n\}}$$

such that $X_n \uparrow X$ for all $\omega \in \Omega$ (see, for instance [60], Theorem 5.1.1).

Definition 83. *(Expectation) For a measurable random variable $X : \Omega \to \mathbb{R} \cup \{+\infty\}$, the expectation value with respect to $\mathbb{P}$ is defined by*

$$\mathbb{E}[X] := \int_{\Omega} X(\omega) d\mathbb{P}(\omega).$$

Whenever X takes the value $+\infty$ on a subset of positive measure, we have $\mathbb{E}[X] = +\infty$.

Note that if $X = \mathbb{1}_A$, we have that $\mathbb{E}[\mathbb{1}_A] = \mathbb{P}(A)$. If X is a simple random variable of the form

$$X = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i},$$

with $|a_i| < \infty$, and $\sum_{i=1}^n A_i = \Omega$, the integral (expectation) of $X \in \mathcal{L}$ with respect is

$$\mathbb{E}[X] = \sum_{i=1}^n a_i \mathbb{P}(A_i).$$

We summarize some properties of expectation in the following:

(i) Non-negativity: If $X \geq 0$, then $\mathbb{E}[X] \geq 0$.

(ii) Linearity: For $c, d \in \mathbb{R}$, we have

$$\mathbb{E}[cX + dY] = c\mathbb{E}[X] + d\mathbb{E}[Y].$$

(iii) Dominated Convergence Theorem: Let $(X_n)_n$, X are random variable. If $X_n \to X$ and there exist a dominating random variable $Y \in L^1$ such that $|X_n| \leq Y$, then

$$\mathbb{E}[X_n] \to \mathbb{E}[X] \quad \text{and} \quad \mathbb{E}[|X_n - X|] \to 0.$$

(iv) Markov Inequality: Suppose X is a non-negative random variable and $X \in L^1$. For any $\alpha > 0$,

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}.$$

L^p-Spaces

Definition 84. *If $p \in [1, +\infty)$, for a measurable random variable $X : \Omega \to \mathbb{R}$, the space $L^p(\Omega, \mathcal{F}, \mathbb{P})$ (or just L^P if there is no ambiguity) is defined by,*

$$L^p(\Omega, \mathcal{F}, \mathbb{P}) := \left\{ X : \Omega \to \mathbb{R} : \ X \textit{ is measurable}, \ \int_\Omega |X(\omega)|^p d\mathbb{P}(\omega) < \infty \right\}. \tag{A.1}$$

In this case, we define the L^p norm of X by

$$\|X\|_p := \left(\int_\Omega |X(\omega)|^p d\mathbb{P}(\omega) \right)^{\frac{1}{p}} = \left(\mathbb{E}[|X|^p] \right)^{\frac{1}{p}}. \tag{A.2}$$

Observe that the space $L^1(\Omega, \mathcal{F}, \mathbb{P})$ is the set of random variables such that $\mathbb{E}[|X|] < +\infty$. If $X \in L^1$, we say that X is integrable with respect to $\mathbb{P}$.

To complete the picture of L^p spaces, we introduce the space of bounded random variables, corresponding to $p = +\infty$, namely.

$$L^\infty(\Omega, \mathcal{F}, \mathbb{P}) := \{ X : \Omega \to \mathbb{R} : \ X \text{ is measurable and there is a } \ C \in \mathbb{R} \text{ such that } |X(\omega)| \leq C, \ \text{a.s. on } \Omega \}, \tag{A.3}$$

which is equipped with the norm

$$\|X\|_\infty := \inf \{ C \geq 0 : |X| \leq C \text{ a.s.} \}.$$

Definition 85. *(Convergence in L^p) Let $p \in [1, \infty]$. We say that a sequence $(X_n)_{n\in\mathbb{N}} \subset L^p$ converges in L^p to $X \in L^p$, written*

$$X_n \xrightarrow{\|\cdot\|_p} X,$$

if

$$\|X_n - X\|_p \to 0, \text{ as } n \to \infty.$$

Definition 86. *(Convergence almost surely) Give a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. A sequence X_n converges almost surely (or converges a.s. for short), denoted*

$$X_n \xrightarrow{a.s.} X,$$

if

$$X_n(\omega) \to X(\omega), \quad \text{for all}, \ \omega \in \Omega \setminus N,$$

where $\mathbb{P}(N) = 0$.

Remark 87. *For $p \in [1, \infty]$, let $(X_n)_{n\in\mathbb{N}}$ be a sequence in L^p such that*

$$X_n \xrightarrow{\|\cdot\|_p} X.$$

Then there exists a subsequence $(X_{n_k})_{k\in\mathbb{N}}$ of $(X_n)_{n\in\mathbb{N}}$, and a function $Y \in L^p$ such that

$$X_{n_k} \xrightarrow{a.s.} X, \quad |X_{n_k}| \leq Y.$$

Theorem 88. *(Scheffe's Lemma) Suppose X_n and X are integrable, and $X_n \xrightarrow{a.s.} X$ Then*

$$X_n \xrightarrow{\|\cdot\|_1} X$$

if and only if

$$\mathbb{E}[|X_n|] \to \mathbb{E}[|X|].$$

We take this opportunity to point out that these limit theorems apply only to sequences, not nets.

Conditional Expectation

Definition 89. *(Conditional expectation) Consider a random variable X defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ for which $\mathbb{E}[|X|] < \infty$. Suppose the σ-algebra $\mathcal{G} \subset \mathcal{F}$. Then there exist a random variable $\mathbb{E}[X|\mathcal{G}]$, called the conditional expectation of X with respect to $\mathcal{G}$, such that*

(i) $\mathbb{E}[X|\mathcal{G}] \in L^1$,

(ii) $\mathbb{E}[X|\mathcal{G}]$ *is $\mathcal{G}$-measurable,*

(iii) For all $A \in \mathcal{G}$, we have

$$\int_A X d\mathbb{P} = \int_A \mathbb{E}[X|\mathcal{G}] d\mathbb{P}.$$

Namely,

$$\mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbb{1}_A], \quad \textit{for all } A \in \mathcal{G}.$$

Conditional expectation inherits many of the properties from ordinary expectation, we itemize some properties of conditional expectation. Note that, all the following formulas are to be understood in an almost sure sense.

(i) Linearity: If $X, Y \in L^1$ and $c, d \in \mathbb{R}$, we have

$$\mathbb{E}[cX + dY|\mathcal{G}] = c\mathbb{E}[X|\mathcal{G}] + d\mathbb{E}[Y|\mathcal{G}].$$

(ii) Monotonicity: If $X, Y \in L^1$ and $X \leq Y$, implies

$$\mathbb{E}[X|\mathcal{G}] \leq \mathbb{E}[Y|\mathcal{G}].$$

In particular, If $X \geq 0$, then $\mathbb{E}[X|\mathcal{G}] \geq 0$.

(iii) Stability: If X is $\mathcal{G}$-measurable; $X \in L^1$ then

$$\mathbb{E}[X|\mathcal{G}] = X.$$

(iv) Product Rule: In general, one can pull out what is known. Let X, Y be random variable satisfying $X, XY \in L^1$. If Y is $\mathcal{G}$-measurable, then

$$\mathbb{E}[XY|\mathcal{G}] = Y\mathbb{E}[X|\mathcal{G}].$$

(v) Law of Total Expectation: If $X \in L^1$, we have

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X].$$

(vi) Conditional Jensen's Inequality: Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a convex function and $\mathbb{E}[|\varphi(X)|] < \infty$. Then for any random variable $X \in L^1$ and σ-algebra $\mathcal{G}$,

$$\varphi(\mathbb{E}[X|\mathcal{G}]) \leq \mathbb{E}[\varphi(X)|\mathcal{G}].$$

(vii) Contraction in $L^p,\ p \geq 1$, then

$$\|E[X|\mathcal{G}]\|_p \leq \|X\|_p.$$

(viii) The Tower Property: Let $\mathcal{H}, \mathcal{G}$ be sub -σ-algebra such that $\mathcal{H} \subset \mathcal{G}$, then

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}].$$

(ix) Dominated Convergence Theorem: If $|X_n| \leq Y$ for all $n \in \mathbb{N}$ where $Y \in L^1$. If $X_n \to X$, then

$$\mathbb{E}[X_n|\mathcal{G}] \to \mathbb{E}[X|\mathcal{G}].$$

Bibliography

[1] B. Acciaio and G. Svindland, "Are law-invariant risk functions concave on distributions?" *Dependence Modeling*, vol. 1, no. 2013, pp. 54–64, 2013.

[2] C. Acerbi and D. Tasche, "On the coherence of expected shortfall," *Journal of Banking & Finance*, vol. 26, no. 7, pp. 1487–1503, 2002.

[3] J. Y. Ahn and N. D. Shyamalkumar, "Asymptotic theory for the empirical Haezendonck–Goovaerts risk measure," *Insurance: Mathematics and Economics*, vol. 55, no. C, pp. 78–90, 2014.

[4] F. Alizadeh and D. Goldfarb, "Second-order cone programming," *Mathematical Programming, Series B*, vol. 95, no. 1, pp. 3–51, 2003.

[5] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, "Coherent measures of risk," *Mathematical finance*, vol. 9, no. 3, pp. 203–228, 1999.

[6] N. Bäuerle and A. Müller, "Stochastic orders and risk measures: consistency and bounds," *Insurance: Mathematics and Economics*, vol. 38, no. 1, pp. 132–148, 2006.

[7] F. Bellini and E. R. Gianin, "On Haezendonck risk measures," *Journal of Banking & Finance*, vol. 32, no. 6, pp. 986–994, 2008.

[8] ——, "Optimal portfolios with Haezendonck risk measures," *Statistics & Risk Modeling*, vol. 26, no. 2, pp. 89–108, 2008.

[9] ——, "Haezendonck–Goovearts risk measures and Orlicz quantiles," *Insurance Mathematics and Economics*, vol. 51, no. 1, pp. 107–114, 2012.

[10] F. Bellini, P. Koch-Medina, C. Munari, and G. Svindland, "Law-invariant functionals on general spaces of random variables," 2020.

[11] S. Biagini and A. Černỳ, "Convex duality and Orlicz spaces in expected utility maximization," *Mathematical Finance*, vol. 30, no. 1, pp. 85–127, 2020.

[12] S. Biagini and M. Frittelli, "On the extension of the Namioka-Klee theorem and on the Fatou property for risk measures," in *Optimality and risk-modern trends in mathematical finance*. Springer, 2009, pp. 1–28.

[13] H. W. Borchers and M. H. W. Borchers, *Package 'pracma'*, 2019.

[14] G. N. Boshnakov, *Package 'cvar'*, 2019.

[15] G. Canna, F. Centrone, and E. Rosazza Gianin, "Haezendonck-goovaerts capital allocation rules," *Available at SSRN 3660385*, 2020.

[16] K. C. B. Charalambos D. Aliprantis, *Infinite Dimensional Analysis: A Hitchhiker's Guide, 3rd edn.* Springer S, 2006.

[17] S. Chen, N. Gao, and F. Xanthos, "The strong Fatou property of risk measures," *Dependence Modeling*, vol. 6, no. 1, pp. 183–196, 2018.

[18] P. Cheridito, F. Delbaen, and M. Kupper, "Coherent and convex monetary risk measures for unbounded cadlag processes," *Finance and Stochastics*, vol. 9, no. 3, pp. 369–387, 2005.

[19] P. Cheridito and T. Li, "Dual characterization of properties of risk measures on Orlicz hearts," *Mathematics and Financial Economics*, vol. 2, no. 1, pp. 29—–55, 2008.

[20] ——, "Risk Measures on Orlicz Hearts," *Mathematical Finance*, vol. 19, no. 2, pp. 189–214, 2009.

[21] A. S. Cherny and P. G. Grigoriev, "Dilatation monotone risk measures are law invariant," *Finance and Stochastics*, vol. 11, no. 2, pp. 291–298, 2007.

[22] R.-A. Dana, "A representation result for concave Schur concave functions," *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, vol. 15, no. 4, pp. 613–634, 2005.

[23] F. Delbaen, "Coherent risk measures," *Blätter der DGVFM*, vol. 24, no. 4, pp. 733–739, 2000.

[24] ——, "Coherent risk measures on general probability spaces," in *Advances in finance and stochastics*. Springer, 2002, pp. 1–37.

[25] D. Dentcheva, S. Penev, and A. Ruszczynski, "Kusuoka representation of higher order dual risk measures," *Annals of Operations Research*, vol. 181, no. 1, pp. 325–335, 2010.

[26] G. A. Edgar and L. Sucheston, *Stopping times and directed processes*. Cambridge University Press, 1992, vol. 47.

[27] D. Filipović and M. Kupper, "Monotone and cash-invariant convex functions and hulls," *Insurance: Mathematics and Economics*, vol. 41, no. 1, pp. 1–16, 2007.

[28] D. Filipovic and G. Svindland, "Convex risk measures on l^p," 2007.

[29] D. Filipović and G. Svindland, "The canonical model space for law-invariant convex risk measures is l^1," *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, vol. 22, no. 3, pp. 585–589, 2012.

[30] H. Föllmer and A. Schied, "Convex measures of risk and trading constraints," *Finance and stochastics*, vol. 6, no. 4, pp. 429–447, 2002.

[31] ——, *Stochastic Finance: An Introduction in Discrete Time*. De Gruyter, 2011.

[32] N. Gao, D. Leung, C. Munari, and F. Xanthos, "Fatou property, representations, and extensions of law-invariant risk measures on general Orlicz spaces," *Finance and Stochastics*, vol. 22, no. 2, pp. 395–415, 2018.

[33] N. Gao, C. Munari, and F. Xanthos, "Stability properties of Haezendonck-Goovaerts premium principles," *Insurance: Mathematics and Economics*, vol. 94, pp. 94 – 99, 2020.

[34] N. Gao and F. Xanthos, "On the C-property and w^*-representations of risk measures." *Mathematical Finance*, vol. 28, no. 2, pp. 748–754, 2018.

[35] M. Goovaerts, R. Kaas, J. Dhaene, and Q. Tang, "Some new classes of consistent risk measures," *Insurance: Mathematics and Economics*, vol. 34, no. 3, pp. 505–516, 2004.

[36] M. Goovaerts, D. Linders, K. Van Weert, and F. Tank, "On the interplay between distortion, mean value and Haezendonck–Goovaerts risk measures," *Insurance: Mathematics and Economics*, vol. 51, no. 1, pp. 10–18, 2012.

[37] M. J. Goovaerts, R. Kaas, J. Dhaene, and Q. Tang, "A unified approach to generate risk measures," *ASTIN Bulletin*, vol. 33, no. 2, pp. 173–191, 2003.

[38] B. Grechuk and M. Zabarankin, "Schur convex functionals: Fatou property and representation," *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, vol. 22, no. 2, pp. 411–418, 2012.

[39] P. G. Grigoriev and J. Leitner, "Dilatation monotone and comonotonic additive risk measures represented as choquet integrals," *Statistics & Decisions*, vol. 24, no. 1/2006, pp. 27–44, 2006.

[40] F. Gómez and Q. Tang, "The gradient allocation principle based on the higher moment risk measure," *Working paper*, 2020.

[41] J. Haezendonck and M. Goovaerts, "A new premium calculation principle based on Orlicz norms," *Insurance: Mathematics and Economics*, vol. 1, no. 1, pp. 41–53, 1982.

[42] N. Horbenko, B. Spangl, S. Desmettre, and E. Massini, *Package 'RobExtremes'*, 2019.

[43] E. Jouini, W. Schachermayer, and N. Touzi, "Law invariant risk measures have the Fatou property," HAL, Post-Print, 2006.

[44] M. Kaina and L. Rüschendorf, "On convex risk measures on l^p-spaces," *Mathematical Methods of Operations Research*, vol. 69, no. 3, pp. 475–495, 2009.

[45] P. Koch-Medina, C. Munari *et al.*, "Law-invariant risk measures: Extension properties and qualitative robustness," *Statistics & Risk Modeling*, vol. 31, no. 3-4, pp. 1–22, 2014.

[46] M. Kohl, P. Ruckdeschel, and M. M. Kohl, *Package 'distrex'*, 2019.

[47] V. Krätschmer, A. Schied, and H. Zähle, "Comparative and qualitative robustness for law-invariant risk measures," *Finance and Stochastics*, vol. 18, no. 2, pp. 271–295, 2014.

[48] P. A. Krokhmal, "Higher moment coherent risk measures," *Quantitative Finance*, vol. 7, no. 4, pp. 373–387, 2007.

[49] S. Kusuoka, "On law invariant coherent risk measures," in *Advances in mathematical economics*. Springer, 2001, pp. 83–95.

[50] J. Leitner, "Balayage monotonous risk measures," *International Journal of Theoretical and Applied Finance (IJTAF)*, vol. 07, no. 07, pp. 887–900, 2004.

[51] F.-B. Liebrich and G. Svindland, "Model spaces for risk measures," *Insurance: Mathematics and Economics*, vol. 77, pp. 150–165, 2017.

[52] ——, “Efficient allocations under law-invariance: A unifying approach,” *Journal of Mathematical Economics*, vol. 84, pp. 28–45, 2019.

[53] Q. Liu, L. Peng, and X. Wang, “Haezendonck–goovaerts risk measure with a heavy tailed loss,” *Insurance: Mathematics and Economics*, vol. 76, no. C, pp. 28–47, 2017.

[54] T. Mao and T. Hu, “Second-order properties of the Haezendonck–Goovaerts risk measure for extreme risks,” *Insurance: Mathematics and Economics*, vol. 51, no. 2, pp. 333–343, 2012.

[55] T. Mao and R. Wang, “Risk aversion in regulatory capital principles,” *SIAM Journal on Financial Mathematics*, vol. 11, no. 1, pp. 169–200, 2020.

[56] Y. Matmoura and S. Penev, “Multistage optimization of option portfolio using higher order coherent risk measures,” *European Journal of Operational Research*, vol. 227, no. 1, pp. 190–198, 2013.

[57] K. Owari, “Maximum lebesgue extension of monotone convex functions,” *Journal of Functional Analysis*, vol. 266, no. 6, pp. 3572–3611, 2014.

[58] F. Peracchi and A. V. Tanase, “On estimating the conditional expected shortfall,” *Applied Stochastic Models in Business and Industry*, vol. 24, no. 5, pp. 471–493, 2008.

[59] M. Rahsepar and F. Xanthos, “On the extension property of dilatation monotone risk measures,” *Statistics & Risk Modeling*, vol. 1, no. ahead-of-print, 2020.

[60] S. I. Resnick, *A Probability Path*. Springer Science & Business Media, 2013.

[61] R. T. Rockafellar and S. Uryasev, “Optimization of conditional value-at-risk,” *Journal of Risk*, vol. 2, pp. 21–41, 2000.

[62] ——, “Conditional value-at-risk for general loss distributions,” *Journal of Banking & Finance*, vol. 26, no. 7, pp. 1443–1471, 2002.

[63] G. Svindland, “Continuity properties of law-invariant (quasi-)convex risk functions on l^∞,” *Mathematics and Financial Economics*, vol. 3, no. 1, pp. 39–43, 2010.

[64] ——, “Dilatation monotonicity and convex order,” *Mathematics and Financial Economics*, vol. 8, no. 3, pp. 241–247, 2014.

[65] Q. Tang and F. Yang, “On the Haezendonck–Goovaerts risk measure for extreme risks,” *Insurance: Mathematics and Economics*, vol. 50, no. 1, pp. 217–227, 2012.

[66] ——, “Extreme value analysis of the Haezendonck–Goovaerts risk measure with a general young function,” *Insurance: Mathematics and Economics*, vol. 59, no. C, pp. 311–320, 2014.

[67] M. Tantrawan and D. H. Leung, “On closedness of law-invariant convex sets in rearrangement invariant spaces,” *Archiv der Mathematik*, vol. 114, no. 2, pp. 175–183, 2020.

[68] H. Wickham and M. H. Wickham, *Package ’ggplot’*, 2020.

[69] A. Zeileis, G. Grothendieck, J. A. Ryan, J. M. Ulrich, F. Andrews, and M. A. Zeileis, *Package ‘zoo’*, 2020.