DEPARTMENT OF MATHEMATICS

Utility-Based Risk Measures and Time Consistency of Dynamic Risk Measures

Sebastian Geissel

Vom Fachbereich Mathematik der Technischen Universität Kaiserslautern
zur Verleihung des akademischen Grades Doktor der Naturwissenschaften
(Doctor rerum naturalium, Dr. rer. nat.) genehmigte Dissertation.

1. Gutachter: Prof. Dr. Jörn Saß, Technische Universität Kaiserslautern
2. Gutachter: Prof. Dr. Stefan Weber, Leibniz Universität Hannover

Datum der Disputation: 6. April 2016

D386

Abstract

This thesis deals with risk measures based on utility functions and time consistency of dynamic risk measures. It is therefore aimed at readers interested in both, the theory of static and dynamic financial risk measures in the sense of Artzner, Delbaen, Eber and Heath [7], [8] and the theory of preferences in the tradition of von Neumann and Morgenstern [134].

A main contribution of this thesis is the introduction of *optimal expected utility* (OEU) risk measures as a new class of utility-based risk measures. We introduce OEU, investigate its main properties, and its applicability to risk measurement and put it in perspective to alternative risk measures and notions of certainty equivalents. To the best of our knowledge, OEU is the only existing utility-based risk measure that is (non-trivial and) coherent if the utility function u has constant relative risk aversion. We present several different risk measures that can be derived with special choices of u and illustrate that OEU reacts in a more sensitive way to slight changes of the probability of a financial loss than value at risk (V@R) and average value at risk.

Further, we propose *implied risk aversion* as a coherent rating methodology for retail structured products (RSPs). Implied risk aversion is based on optimal expected utility risk measures and, in contrast to standard V@R-based ratings, takes into account both the upside potential and the downside risks of such products. In addition, implied risk aversion is easily interpreted in terms of an individual investor's risk aversion: A product is attractive (unattractive) for an investor if its implied risk aversion is higher (lower) than his individual risk aversion. We illustrate this approach in a case study with more than 15,000 warrants on DAX® and find that implied risk aversion is able to identify favorable products; in particular, implied risk aversion is not necessarily increasing with respect to the strikes of call warrants.

Another main focus of this thesis is on consistency of dynamic risk measures. To this end, we study risk measures on the space of distributions, discuss concavity on the level of distributions and slightly generalize Weber's [137] findings on the relation of time consistent dynamic risk measures to static risk measures to the case of dynamic risk measures with time-dependent parameters. Finally, this thesis investigates how recursively composed dynamic risk measures in discrete time, which are time consistent by construction, can be related to corresponding dynamic risk measures in continuous time. We present different approaches to establish this link and outline the theoretical basis and the practical benefits of this relation. The thesis concludes with a numerical implementation of this theory.

Zusammenfassung

Diese Arbeit beschäftigt sich mit nutzenbasierten Risikomaßen und mit Zeitkonsistenz dynamischer Risikomaße. Folglich richtet sie sich sowohl an Leser mit Interesse an der Theorie statischer und dynamischer Risikomaße im Sinne von Artzner, Delbaen, Eber und Heath [7], [8] und an der Nutzentheorie nach von Neumann und Morgenstern [134].

Ein wesentlicher Beitrag dieser Arbeit ist die Einführung der *optimalen Erwartungsnutzen-Risikomaße* (OEU) als eine neue Klasse nutzenbasierter Risikomaße. Wir führen OEU ein, ermitteln ihre wichtigsten Eigenschaften und ihre Eignung für die Risikobewertung und vergleichen sie mit anderen Risikomaßen und Sicherheitsäquivalenten. Nach unserer Kenntnis sind OEU die einzig existierenden nutzenbasierten Risikomaße für Nutzenfunktionen u mit konstanter relativer Risikoaversion, die (nicht trivial und) koheränt sind. Wir präsentieren verschiedene Risikomaße, die sich für eine geeignete Wahl von u ergeben und veranschaulichen, dass OEU sensibler auf geringfügige Änderungen der Wahrscheinlichkeit für finanziellen Verlust reagieren als Value at Risk (V@R) und Average Value at Risk.

Außerdem schlagen wir das Konzept der *implizierten Risikoaversion* als eine koheränte Bewertungsmethode für Zertifikate vor. Die implizierte Risikoaversion basiert auf OEU und berücksichtigt - im Gegensatz zu den V@R-basierten Bewertungen - sowohl die Ertragschancen als auch die Verlustrisiken dieser Produkte. Zudem lässt sich implizierte Risikoaversion gut hinsichtlich der Risikoaversion eines Privatanlegers interpretieren: Ein Produkt erscheint einem Anleger (un)attraktiv, wenn dessen implizierte Risikoaversion (kleiner) größer ist als seine persönliche Risikoaversion. Wir veranschaulichen dies mit einer Fallstudie von mehr als 15,000 DAX®-Optionsscheinen und zeigen auf, dass implizierte Risikoaversion vorteilhafte Produkte bestimmen kann. Insbesondere steigt die implizierte Risikoaversion nicht generell in Bezug auf den Ausübungspreis von Call-Optionsscheinen.

Ein weiterer Schwerpunkt der Arbeit liegt auf der Konsistenz von dynamischen Risikomaßen. Hierzu untersuchen wir Risikomaße auf dem Raum der Wahrscheinlichkeitsverteilungen, erörtern Konkavität hinsichtlich Verteilungen und verallgemeinern die Erkenntnisse von Weber [137] über den Zusammenhang von zeitkonsistenten dynamischen Risikomaßen und statischen Risikomaßen geringfügig für den Fall von dynamischen Risikomaßen mit zeitabhängigen Parametern. Abschließend untersuchen wir, wie rekursiv zusammengesetzte dynamische Risikomaße in diskreter Zeit, die per Konstruktion zeitkonsistent sind, mit entsprechenden dynamischen Risikomaßen in stetiger Zeit in Verbindung gebracht werden können. Wir stellen verschiedene Ansätze vor, wie eine solche Verbindung hergestellt werden kann, und stellen das theoretische Fundament und den praktischen Nutzen dieses Zusammenhangs dar. Die Arbeit endet mit einer numerischen Umsetzung dieser Theorie.

Acknowledgments

First and foremost, I sincerely thank my supervisor Prof. Dr. Jörn Saß for his support, his constant interest in my work, and for his helpful and friendly mentoring during the past years. I have learned a lot from him and appreciate his availability and the fruitful discussions during the work on this thesis. I am also grateful to Prof. Dr. Stefan Weber for accepting to act as a referee for this thesis.

Special thanks are due to Prof. Dr. Frank Seifried. His encouraging and precious suggestions have been of great value for me. My appreciation also goes to my colleagues and friends of the Financial Mathematics Group at the University of Kaiserslautern and to Dr. Holger Fink from Ludwig-Maximilians-Universität Munich.

I would like to thank Maria, Rudi, Viola, Kader and my friends from university and from my hometown for their continued support and encouragement. I really appreciate having had the privilege of studying without any serious worries.

I gratefully acknowledge financial and ideal support by the Stiftung der Deutschen Wirtschaft.

Contents

List of Figures

List of Tables

List of Listings

Preliminaries

Notation

Throughout this thesis we make use of the following notations:

(i) We say that a property holds *strictly* if it is valid without exception. For example, a real number x is *positive* if $x \geq 0$ and *strictly positive* if $x > 0$.

(ii) Properties involving random variables on a probability space $(\Omega,\mathcal{F},P)$ are always to be understood in a P-almost sure (a.s.) sense without explicitly saying so. Saying that a random variable X is, e.g., *positive* means that $X \geq 0$ P-a.s.

(iii) $\mathbf{1}_A$ denotes the *indicator function* of A, i.e.,

$$\mathbf{1}_A(x) := \begin{cases} 1, & x \in A, \\ 0, & x \notin A \end{cases} .$$

(iv) We set $\mathbb{R}_+ := (0,\infty)$.

(v) We denote the minimum of two real numbers by $x \wedge y := \min\{x,y\}$.

For the sake of readability, the masculine form is used throughout this thesis. Wherever appropriate, use of the feminine form is, of course, implicit.

CHAPTER 1

Introduction

Financial mathematics aims at modeling the financial market and financial products in order to make theoretical calculations possible. This allows, among other things, assessing financial risks and finding optimal investment under a given theoretical framework. In this thesis we address questions assigned to utility-based approaches to calculating financial risks and to risk measurement in a dynamic time setting.

Measuring financial risks

Financial risk management includes a large number of different types of risks which can be classified in the following categories:

- *Market risk*: Risk that the value of an investment will decrease due to unexpected changes in underlying components such as prices or rates, or risk of changing volatility of market prices
- *Credit risk*: Risk of a debtor of not receiving required repayments on outstanding loans or bonds due to a default of the borrower, or risk of losing market value due to changes in the credit quality of counterparties
- *Interest rate risk*: Risk for bond owners from fluctuating interest rates depending on the sensitivity of the respective bond(s) to interest rate changes
- *Liquidity risk*: Risk of an unexpected rise in the costs of adjusting investments (widened bid-ask spreads), or risk of losing access to refinancing at moments of financial distress
- *Model risk*: Risk of using a misspecified model for evaluating risk
- Other types of risk such as *operational risk*, *systemic risk* or *legal risk*.

In order to identify financial risks, that is associating to each element of a set of identified risky elements a real number, we must first agree on a theoretical framework for the nature and form of appearance of risk. In most parts of this thesis we consider random variables to reflect financial risks, but, under a suitable condition of law-invariance, we also represent risk by probability distributions. Either way, summarizing (risk-)relevant information about a financial position in a single number is a key task for investors, regulators and financial institutions and a necessary prerequisite to take financial decisions under uncertainty; see Delbaen [45].

The essential idea of risk measures is to quantify risk as a capital requirement. More precisely, risk measures have been developed to determine minimum capital reserves that have to be deposited by financial institutions in order to ensure their financial stability. This is clearly a very delicate matter and has to be done as accurately as possible. Incorrect measurements of risk can either lead to opportunity costs if a financial institution puts aside too much safety capital or to threat scenarios if too little safety capital is deposited for potential risks arising. Therefore the theory of financial risk measures is of particular interest to both, financial regulatory authorities that try to maintain the stability of the financial system, and to financial institutions that are interested in having a well functioning risk controlling system.

In the seminal work of Markowitz [104], the variance of a random variable was proposed as a measure of its risk. Other approaches to evaluate financial risks include the concepts of expected losses, certainty equivalents, deviation measures, see Rockafellar, Uryasev and Zabarankin [119], and quantile-based measures, see Dowd and Blake [53]. In this thesis, however, we exclusively focus on the axiomatic analysis of risk through *risk measures* as initiated by Artzner, Delbaen, Eber and Heath [7], [8].

Risk measures on the space of random variables

The most common assumption on risk measures in the literature is to model financial positions as random variables X which have an uncertain outcome. In the fundamental papers of Artzner, Delbaen, Eber and Heath [7], [8], which initiated an axiomatic analysis of risk assessment, some desirable properties of risk measures were formulated: A riskier position requires more reserve capital; a cash reserve reduces the capital required by its present value; a merger does not imply extra risk; and position size does not influence risk.

The evaluation of the riskiness of a financial position from today's point of view for a fixed time horizon is called static risk measurement. Think of, for example, measuring the risk of a stock position in 10 days. In the light of the usually long-term business environments where such risk measures are often applied, this methodology raises questions: How can we consider future incoming information on the market environment and if we update risks, how are risks measured at different points in time interrelated? Dealing with available future information on considered financial positions and the consideration of meaningful interconnections between risks measured in different time periods eventually opens up the theory of dynamic risk measurement.

In Chapter 2 we give an overview of the theory of risk measures on the space of random variables in a static and a dynamic framework and review standard examples of risk measures including value at risk, average value at risk and the entropic risk measure.

A common view on risk measures is that they represent the minimal cash amount that has to be added to X to make it *acceptable*. Hence, the economic interpretation of acceptability plays a crucial role in this context. Due to the fact that often financial products bear a higher risk the larger the potential gains of the product, the definition of acceptable financial positions can be the result of complex and controversial discussion. Therefore we decided to consider economic preference theory and eventually come up with a new class of risk measures that can include an investor's risk attitude.

Risk measures based on utility functions

The economic theory of preferences comprises the concept of utility functions u which formalize the attitude towards risk of an exemplary investor in the financial market. For instance, an investor is risk averse in the sense that he dislikes all zero-mean risk at all levels of wealth if and only if his utility function is concave; see, among others, Gollier [81, Proposition 6]. Risk aversion can then be quantified by the *certainty equivalent*, defined as $\mathbb{C}_u[X] := u^{-1}\left(\mathbb{E}\left[u(X)\right]\right)$, which can be understood as the sure amount that makes an investor indifferent between investing in X and receiving $\mathbb{C}_u[X]$.

Risk measures based on utility functions combine the theory of risk measures and classical topics of economic preference theory. For example, Föllmer and Schied [69, Example 4.13] and Müller [108] consider the negative of a classical certainty equivalent, $-u^{-1}\left(\mathbb{E}\left[u\left(X\right)\right]\right)$, as a risk measure. Other constructions of risk measures based on utility functions have been introduced by Ben-Tal and Teboulle [18], [19], Föllmer and Schied [67] and Krokhmal [97], among others. We give a more detailed introduction into this subject in Chapter 3.

In Chapter 4 we introduce, investigate and exemplify the new class of *optimal expected utility* (OEU) risk measures that are generated by utility functions via an associated optimal investment problem. A significant advantage of OEU over other utility-based risk measures is that it allows for most commonly used specifications of utility functions (in particular, power, logarithmic and exponential utilities), and that it is (non-trivial and) coherent if the utility function u has constant relative risk aversion

Evaluating risk and return of retail structured products

As mentioned before, risk measures are also well suited for evaluating future payoffs of financial products which make them of interest in the context of classifying retail structured products (RSPs) and thereby backing up investment decisions in these products. Since RSPs can lead to significant losses on the one hand, while alluring with potential gains on the other hand, we conclude that risk measures which enable practitioners to take into consideration the whole payoff structure of financial positions, are of particular significance for the aforementioned purpose.

Chapter 5 includes a practical proposal of an alternative rating system for RSPs based on their respective *implied risk aversion*. Briefly, the implied risk aversion of a financial payoff X is the parameter $\gamma_0(X)$ such that an investor with risk aversion $\gamma_0(X)$ is indifferent between the financial position X and a zero investment.

Risk measures on the space of probability distributions

In Chapter 6 we deal with law-invariant risk measures which only depend on the distribution of financial positions under the given probability measure. These risk measures can also be interpreted as functionals on the space of distributions which are usually referred to as lotteries in decision theory. On the basis of Weber and Schmidt [138] and Acciaio and Svindland [2], we discuss certain properties of risk measures on the space of distributions. Hereby, the relation of the risk evaluation of *mixture distributions*, which are convex combinations of distributions, to the risks of the consisting distributions is of particular interest. We also consider an extreme event scenario in order to compare between value at risk, average value at risk and optimal expected utility risk measures with respect to their sensitivity to potentially large losses.

Weber [137] established a relation between attributes of static risk measures on the space of distributions and consistency properties of corresponding dynamic risk measures on the space of random variables. In view of Weber's findings, we study consistency properties of dynamic risk measures with time-dependent parameters.

Dynamic risk measures in a discrete and a continuous time setting

For the sake of controlling the riskiness of a financial position X at different times or even at all times between today and a given maturity time T, we need *dynamic risk measures*. The central idea of dynamic risk measures is to evaluate the riskiness of X at T from the standpoint of information which is available at time $t \leq T$. The natural starting point for a dynamic risk evaluation is $t = T$ where the payoff of X is "known" in the sense that it is, for example, the result of a simulation. Then, we can recursively move backwards in time and measure the risk of X at increasingly early points in time. Thereby we not only get the initial ($t = 0$) risk of X, but we also receive intermediate results which may serve as an indication of the amount of necessary capital reserves at future times $t \in [0,T]$. Following this idea, one can construct time consistent risk measures in discrete time by recursively composing one-period risk measures. In a continuous time setting, however, almost any coherent/convex dynamic risk measure comes from a conditional g-expectation which is the solution of a backward stochastic differential equation; see Rosazza Gianin [121], among others.

In Chapter 7 we present examples of composed time consistent dynamic risk measures in discrete time and show the necessity and a possible way of rescaling the one-period risk measures. For composed (scaled) versions of popular risk measures such as value at risk, average value at risk and the entropic risk measure, we work out drivers of corresponding BSDEs and compare the resulting dynamic risk measures in continuous time to the related composed discrete-time dynamic risk measures when the size of the discrete time grid goes to zero. Our results allow for a nice interpretation of dynamic risk measures as, on the one hand we understand the behavior of discrete-time compositions in any fixed time interval, and on the other hand [121] provides a nice interpretation of the functionals g which are crucial for the construction of the continuous-time dynamic risk measures.

CHAPTER 2

Risk measures on the space of random variables

Risk measures have been developed in order to determine minimum capital reserves that have to be deposited by financial institutions to ensure their financial stability. Other applications of risk measures include the evaluation of future losses, the acceptability of risk exposures as well as pricing. The landmark articles of Artzner, Delbaen, Eber and Heath [7, 8] initiated a rigorous, axiomatic analysis of risk assessments, and also introduced the notion of coherent risk measures. Subsequently and independently of each other, Föllmer and Schied [67], Fritelli and Rosazza Gianin [74], Heath [83] and Heath and Ku [84] replaced the subadditivity and positive homogeneity properties by a convexity condition and thus established the more general concept of convex risk measures. Since then, convex and coherent risk measures have been intensively studied and extended in various directions; see, for example, Acerbi and Tasche [4], Delbaen [44], Detlefsen and Scandolo [46], Rockafellar and Uryasev [117, 118], and the references therein. For an overview of the theory of coherent and convex risk measures, we refer to Delbaen [43] and Föllmer and Schied [69, Chapter 4].

This chapter includes preliminaries and known results on risk measures on the space of random variables that are necessary for our further work: In Section 2.1 we introduce the theoretical setup and some notations. Section 2.2 provides main definitions and results of static risk measures which is then extended to conditional risk measures in Section 2.3 and finally to dynamic risk measures in Section 2.4. Section 2.5 concludes with some results on dynamic risk measures in a continuous time setting.

2.1 The setup and notations

We fix a filtration $(\mathcal{F}_t)_{0\leq t\leq T}$, $T > 0$, on the probability space $(\Omega,\mathcal{F},P)$ such that $\mathcal{F}_0 = \{\varnothing,\Omega\}$ and $\mathcal{F}_T = \mathcal{F}$. A *financial position* is a random variable $X : \Omega \to \mathbb{R}$ defined on $(\Omega, \mathcal{F}, P)$. By a standard convention, see, e.g., [8], $X(\omega)$ represents the position's value, net of all costs including those to set up the position, at the end of the investment period if the scenario $\omega \in \Omega$ materializes. We take the space $\mathcal{X} := L^\infty(\Omega,\mathcal{F},P)$ as the set of all terminal values of financial positions and define the subspace $L^\infty(\mathcal{F}_t) := L^\infty(\Omega,\mathcal{F}_t,P)$ as the set of all essentially bounded $\mathcal{F}_t$-measurable random variables. In the following we briefly write

$$X_{\min} := \operatorname{ess\,inf} X$$

for the left support of $X \in \mathcal{X}$. Finally, we denote by $\beta > 0$ the *risk-free discount factor* in the financial market; thus, if p denotes the market price of a default-free zero-coupon bond and T is the relevant time horizon, then $\beta = p(0,T)$. Note that although $\beta \le 1$ seems a realistic assumption on the risk-free discount factor, we formally also allow for $\beta > 1$ corresponding to negative interest rates.

2.2 Static risk measures

At this point we recall some of the most important aspects and results of the theory of static risk measures. We refer to [69, Chapter 4] for a widespread presentation of this concept.

2.2.1 Basic definitions and main properties of static risk measures

The axiomatic approach to risk measures sets a number of desirable characteristics of risk measures which form the basis for the following definition.

2.1 Definition. A map $\rho : \mathcal{X} \to \mathbb{R}$ is called a *risk measure* if it satisfies, for all $X,Y \in \mathcal{X}$,

(M) *Monotonicity*: $\rho(X) \ge \rho(Y)$ if $X \le Y$,

(CI) *Cash invariance*: $\rho(X+m) = \rho(X) - \beta m$ for all $m \in \mathbb{R}$,

(N) *Normalization*: $\rho(0) = 0$.

A risk measure ρ is called *convex* if it satisfies

(C) *Convexity*: $\rho(\lambda X + (1-\lambda)Y) \le \lambda\rho(X) + (1-\lambda)\rho(Y)$ for every $\lambda \in [0,1]$.

A convex risk measure ρ is called *coherent* if it additionally satisfies

(PH) *Positive homogeneity*: $\rho(\lambda X) = \lambda\rho(X)$ for any $\lambda \ge 0$.

If we adopt the traditional point of view that $\rho(X)$ represents the capital required to support a risky position $X \in \mathcal{X}$, then monotonicity, cash invariance and normalization are intuitive and natural properties that codify the following properties: The downside risk of the position Y is smaller than the risk of position X if the values of a position Y are always better than those of a position X; a cash reserve reduces the capital required by its present value; and a zero position is riskless. In particular, cash-invariance implies that $\rho(X)$ is just the amount of money that can be added to the position X in order to make it acceptable; see Chapter 5 where the concept of implied risk aversion is based on this interpretation. Convexity implies that the risk measure supports diversification, and coherence means that the position's risk is proportional to its size. Under the assumption of positive homogeneity, (C) is equivalent to

(S) *Subadditivity*: $\rho(X+Y) \le \rho(X) + \rho(Y)$,

see Remark A.1, which means that the risk of two positions together cannot be worse than adding the risks of the positions separately. This property can be seen as the mathematical counterpart of the idea that diversification reduces risk (*diversification principle*). However, there is a controversial debate about the economic sense of (PH), in which critics claim that positively homogeneous risk measures do not penalize concentration of risk and associated liquidity problems; see, among others, Daníelsson, Jorgensen, Samorodnitsky, Sarma and de Vries [41], Dhaene, Goovaerts and Kaas [48], Dhaene, Laeven, Vanduffel, Darkiewicz and Goovaerts [49] and Kou, Peng and Heyde [95].

2.2 Remark. In our formulation of (CI), we follow Artzner, Delbaen, Eber and Heath [8] and work with actual, rather than discounted, payoffs. In their notation,

$$\rho\left(X + (1+r)m\right) = \rho(X) - m \quad \text{for } m \in \mathbb{R},$$

where $\beta = 1/(1+r)$ and r is the interest rate of a (risk-free) reference instrument. In much of the relevant literature, one works with discounted payoffs; this relates to our definition via

$$\bar{\rho}(\beta X) = \rho(X)$$

and in that case (CI) reads

$$\bar{\rho}(\bar{X} + \bar{m}) = \bar{\rho}(\bar{X}) - \bar{m} \quad \text{for } \bar{m} \in \mathbb{R}.$$

These two formulations are, of course, equivalent; if not otherwise indicated, in all that follows, we use actual, un-discounted, payoffs and transfer existing definitions into that notation without further mentioning.

The classes of all positions $X \in \mathcal{X}$ which do (not) require additional risk capital are defined as follows:

2.3 Definition. The *acceptance set* of a risk measure ρ includes all positions which are acceptable in the sense that they have negative risk, i.e., they do not require additional capital:

$$\mathcal{N}_\rho := \{X \in \mathcal{X} \mid \rho(X) \le 0\}.$$

The set of all positions with strictly positive risk

$$\mathcal{N}_\rho^c := \{X \in \mathcal{X} \mid \rho(X) > 0\}.$$

is called the *rejection set* of ρ.

We now present examples of well-known risk measures which are calculated from a quantile of potential losses of X.

2.4 Example. (a) The *value at risk* (V@R) at level $\lambda \in (0,1)$ of a position $X \in \mathcal{X}$ is defined by

$$V@R^\lambda(X) = \inf\{m \in \mathbb{R} : P(m + \beta X < 0) \le \lambda\}.$$

The probability that the loss of X exceeds $V@R^\lambda(X)$ is less than or equal to λ.

(b) The *average value at risk* (AV@R) at level $\lambda \in (0,1]$ of a position $X \in \mathcal{X}$ is defined as

$$AV@R^{\lambda}(X) = \frac{1}{\lambda}\int_0^{\lambda} V@R^{\gamma}(X)d\gamma.$$

Average value at risk is, in simple terms, the expected value of the losses beyond the V@R point. It is also called *conditional value at risk* or *expected shortfall.*

2.5 Remark. (a) $V@R$ is positively homogeneous, see Remark A.2, but in general it is not subadditive and can therefore discourage diversification; see Remark A.3.

(b) $AV@R$ is subadditive, see Acerbi, Nordio and Sirtori [3, Appendix], and also satisfies all other axioms of a coherent risk measure.

(c) For continuously distributed X, $AV@R^{\lambda}(X)$ coincides with tail conditional expectation,

$$TCE^{\lambda}(X) := \mathbb{E}[-X : -X \geq V@R^{\lambda}(X)];$$

see [69, Corollary 4.54]. By means of the definition of TCE, we can easily prove that AV@R dominates V@R: For any $X \in \mathcal{X}$,

$$\begin{aligned} AV@R^{\lambda}(X) &\geq TCE^{\lambda}(X) \\ &= \mathbb{E}[-X : -X \geq V@R^{\lambda}(X)] \\ &\geq V@R^{\lambda}(X). \end{aligned}$$

Average value at risk is even the smallest law-invariant convex risk measure which is continuous from above that dominates value at risk; see [69, Theorem 4.67].

(d) However, despite their immense popularity, both, value at risk and average value at risk, have drawn much criticism from academia and industry. V@R does not generally support portfolio diversification which contrasts to the axiomatic properties as formulated in Definition 2.1. Moreover, value at risk does not take potential losses beyond the chosen level λ into risk assessment. One might, however, argue that it is far more important to worry about the cases when losses exceed V@R. Thus extreme loss scenarios might be overseen by V@R and we can think of portfolios which have an identical V@R and hence appear equally risky to decision-makers, but have considerably different loss scenarios. Also, as mentioned by Delbaen [45], when used inside a financial institution, value at risk favors the practice “take the money and run”: It encourages traders to take positions, where in, say 99% of the cases, the trader gets a gain but there are extreme loss scenarios that may cause bankruptcy of the institution which occurs with a probability of 1%, i.e., they are hidden beyond the V@R level λ (in case $\lambda > 1\%$). Average value at risk on the other hand includes any possible losses beyond λ, but, since the level is often chosen very low (5% or 1%), it usually requires a high number of data to ensure an accurate calculation of AV@R. This drawback is mentioned in Bellinni and Bignozzi [17] and in Daníelsson [40], among others. Since V@R must always be calculated to get AV@R, we point out that these risk measures may not be seen as alternative choice to each other but that average value at risk may rather be understood as an extension to the concept of value at risk.

Another approach to assessing risk is including utility functions as done in the following example:

2.6 Example. For $X \in \mathcal{X}$, we define the *entropic risk measure* as

$$\rho^{\text{ent}}(X) = \frac{1}{\gamma} \ln \mathbb{E}\left[\exp(-\gamma\beta X)\right],$$

where $\gamma \in \mathbb{R}_+$ is the risk aversion parameter.

The entropic risk measure is introduced in [67] and [74]; for further information on ρ^{ent} we refer to Föllmer and Knispel [65] and the references therein. We refer to Chapter 3 for a more detailed introduction to utility-based risk measures.

2.2.2 Robust representation of static risk measures

By $\mathcal{M}_1(P) := \mathcal{M}_1(\Omega,\mathcal{F},P)$ we denote the set of all probability measures on $(\Omega,\mathcal{F})$ that are absolutely continuous with respect to P. The following theorem characterizes convex risk measures on $\mathcal{X}$ that can be represented by a penalty function concentrated on $\mathcal{M}_1(P)$; it corresponds to [69, Theorem 4.33 and Corollary 4.37].

2.7 Theorem. *Suppose $\rho : L^\infty \to \mathbb{R}$ is a convex risk measure. Then the following conditions are equivalent:*

(i) ρ can be represented by some penalty function on $\mathcal{M}_1(P)$.

(ii) ρ can be represented by the restriction of the minimal penalty function $\vartheta_{\min}$ to $\mathcal{M}_1(P)$

$$\rho(X) = \sup_{Q \in \mathcal{M}_1(P)} \left(\mathbb{E}_Q[-\beta X] - \vartheta_{\min}(Q)\right), \quad X \in L^\infty. \tag{2.1}$$

(iii) ρ is continuous from above: If $X_n \searrow X$ then $\rho(X_n) \nearrow \rho(X)$.

(iv) ρ has the following Fatou property: for any bounded sequence (X_n) which converges to some X,

$$\rho(X) \le \liminf_{n \uparrow \infty} \rho(X_n).$$

(v) ρ is lower semi-continuous for the weak topology $\sigma(L^\infty,L^1)$.*

(vi) The acceptance set $\mathcal{N}_\rho$ of ρ is weak closed in L^∞, i.e., $\mathcal{N}_\rho$ is closed with respect to the topology $\sigma(L^\infty,L^1)$.*

Moreover, under these conditions, ρ is coherent if and only if $\vartheta_{\min}$ takes only the values 0 and ∞. In this case, (2.1) becomes

$$\rho(X) = \sup_{Q \in \mathcal{Q}^{\max}} \mathbb{E}_Q[-\beta X], \quad X \in L^\infty,$$

where $\mathcal{Q}^{\max} := \{Q \in \mathcal{M}_1(P) : \vartheta_{\min}(Q) = 0\}$.

In (2.1) we can interpret the elements of $\mathcal{M}_1(P)$ as possible probabilistic models which are taken more or less seriously depending on the outcome of $\vartheta_{\min}(Q)$; then one takes the worst penalized expected loss over the class $\mathcal{M}_1(P)$. This interpretation is perfectly suitable for illustrating model risk; see, for example, Föllmer and Schied [68, page 6].

2.8 Remark. The entropic risk measure as introduced in Example 2.6 is continuous from above and can therefore be represented as in (2.1) by its minimal penalty function ϑ^{ent}. Due to standard duality results, ϑ^{ent} is given by $\vartheta^{\text{ent}}(Q) := \frac{1}{\gamma} H(Q \mid P)$, where

$$H(Q \mid P) := \sup_{X \in L^\infty} \left(\mathbb{E}_Q[-\beta X] - \ln \mathbb{E}[\exp(-\beta X)] \right)$$

is the relative entropy of $Q \ll P$; see, e.g., [69, Section 3.2, Section 4.9].

2.3 Conditional risk measures

The risk measures mentioned so far are one-period risk measures. However, in practice, financial risks have a dynamic component in two respects: Firstly, in a multi-period framework, there is uncertainty about the time value of money and cash-flows can occur such that the financial position itself is dynamic. In this respect, stochastic processes are better suited for modeling financial positions than random variables in a dynamic setting. Risk measures on processes, however, are not considered in this thesis; we refer to the works of Cheridito, Delbaen and Kupper [32], [33], [34], Cheridito and Kupper [35], Föllmer and Penner [66], Fritelli and Scandolo [76] and Riedel [116] instead.

Due to a constant information flow (e.g., on the market) we want to work with risk measures that are updated in a multi-period framework. In such a dynamic setting, a conditional risk measure ρ_t assigns to each terminal payoff X an $\mathcal{F}_t$-measurable random variable $\rho_t(X)$ that quantifies the risk of the position X conditional on $\mathcal{F}_t$, the information available (e.g., to investors or financial institutions) at time t. In this sense it can be seen as a natural extension to the concept of static risk measurement.

2.3.1 Basic definitions and main properties of conditional risk measures

In the following we introduce dynamic risk measures for terminal payoffs.

2.9 Definition. For $0 \le t \le T$ a map $\rho_t : \mathcal{X} \to L^\infty(\mathcal{F}_t)$ is called a *conditional risk measure* if it satisfies the following properties for all $X, Y \in \mathcal{X}$:

(M) *Monotonicity* : $\rho_t(X) \ge \rho_t(Y)$ if $X \le Y$.

(CCI) *Conditional cash-invariance* : $\rho_t(X + m) = \rho_t(X) - \beta_t^T m$ for all $m \in L^\infty(\mathcal{F}_t)$.

(N) *Normalization* : $\rho_t(0) = 0$.

By the $\mathcal{F}_t$-measurable random variable β_t^T we denote the value of a default-free bond at time t with face value 1 and maturity time T. A conditional risk measure ρ_t is called *convex* if it additionally satisfies:

(CC) *Conditional convexity* : $\rho_t(\lambda X + (1-\lambda)Y) \le \lambda \rho_t(X) + (1-\lambda)\rho_t(Y)$, for $\lambda \in L^\infty(\mathcal{F}_t)$

for every $\lambda \in [0,1]$.

A conditional convex risk measure ρ_t is called *coherent* if it additionally satisfies

(CPH) *Conditional positive homogeneity* : $\rho_t(\lambda X) = \lambda \rho_t(X)$, for $\lambda \in L^\infty(\mathcal{F}_t)$ for any $\lambda \geq 0$.

The economic interpretation of the properties of characterizing conditional risk measures are the same as in the static case. In particular, due to conditional translation invariance, a conditional risk measure determines a conditional capital requirement.

2.10 Definition. The *acceptance set* of a conditional risk measure ρ_t is defined as:

$$\mathcal{N}_t := \{X \in L^\infty \,:\, \rho_t(X) \leq 0\},$$

and the *rejection set* of ρ_t is

$$\mathcal{N}_t^c := \{X \in L^\infty \,:\, \rho_t(X) > 0\}.$$

Note that a conditional risk measure ρ_t is uniquely determined by its acceptance set since

$$\rho_t(X) = \text{ess inf}\{Y \in L^\infty(\mathcal{F}_t) \,:\, X + Y \in \mathcal{A}_t\}.$$

2.11 Example. (a) For any $\lambda \in (0,1]$, the *conditional value at risk at level* λ is defined as

$$V@R_t^{\lambda_t}(X) = \text{ess inf}\{m_t \in L^\infty(\mathcal{F}_t) \,:\, P(\beta_t^T X + m_t < 0 \mid \mathcal{F}_t) \leq \lambda_t\},$$

where $P(A \mid \mathcal{F}_t)$ is the conditional expectation $\mathbb{E}[\mathbf{1}_A \mid \mathcal{F}_t]$.

(b) For $\lambda_t \in L^\infty(\mathcal{F}_t), 0 < \lambda_t \leq 1$ let $\mathcal{Q}_t^{\lambda_t}$ denote the set of all measures $Q \in \mathcal{P}_t := \{Q \in \mathcal{Q}_t \,:\, Q = P|_{\mathcal{F}_t}\}$, with

$$\mathcal{Q}_t := \{Q \in \mathcal{M}_1(P) \,:\, Q = P|_{\mathcal{F}_t}\},$$

whose density dQ/dP is P-a.s. bounded by $1/\lambda_t$. The resulting coherent conditional risk measure

$$AV@R_t^{\lambda_t}(X) = \underset{Q \in \mathcal{Q}_t^{\lambda_t}}{\text{ess sup}}\, \mathbb{E}_Q[-\beta_t^T X \mid \mathcal{F}_t]$$

is called *conditional average value at risk at level* λ_t.

(c) For a risk aversion parameter $\gamma \in \mathbb{R}_+$,

$$\rho_{t,\gamma}^{\text{ent}}(X) = \frac{1}{\gamma} \ln \mathbb{E}[\exp(-\gamma \beta_t^T X) \mid \mathcal{F}_t], \quad X \in \mathcal{X},$$

is called the *conditional entropic risk measure.*

2.4 Dynamic risk measures and time consistency

Dynamic risk measures are defined as a collection of conditional risk measures adapted to the underlying filtration. We denote by I the set of points in time at which we consider conditional risk measures and distinguish between dynamic risk measures in discrete time, $I = \{t_0 = 0, t_1, ..., t_N = T\}$, and in continuous time, $I = [0,T]$. Given such a sequence $(\rho_t)_{t\in I}$ it is an obvious question to ask how risk evaluation between different points in time t and s between today and the maturity time of the financial product T are related to each other. Different definitions of consistency of dynamic risk measures are given in order to answer this question. In this section we give basic definitions, remarks and some selected results on dynamic risk measures which we use in the course of this thesis. For a more detailed introduction to this topic, we particularly recommend the review article [1] and the references therein. We refer to Chapter 7 for a further examination of consistency of dynamic risk measures.

2.12 Definition. We call a collection of mappings $(\rho_t)_{t\in I}$, $\rho_t : \mathcal{X} \to L^\infty(\mathcal{F}_t)$ a *dynamic risk measure* if every ρ_t is a conditional risk measure.

We call $(\rho_t)_{t\in I}$ a dynamic risk measure in continuous time (CDRM) if $I = [0,T]$. $(\rho_t)_{t\in I}$ is called a dynamic risk measure in discrete time (DDRM) if $I = \{t_0,...,t_N\}$, where $0 = t_0 < t_1 < ... < t_N = T$. Note that T can also be ∞.

For any $t \in I$ and for any $X \in \mathcal{X}$, $\rho_t(X)$ is a random variable, therefore for any fixed t any equality and inequality of ρ_t should be understood to be valid P-a.s. Note that the "boundary condition" $\rho_T(X) = -X$ for all $X \in \mathcal{X}$ follows directly from the definition of the acceptance set of a conditional risk measure ρ_t in Definition 2.10.

2.13 Definition. A dynamic risk measure $(\rho_t)_{t\in I}$ is called *convex* if every ρ_t satisfies conditional convexity (CC). A dynamic convex risk measure $(\rho_t)_{t\in I}$ is called *coherent* if every ρ_t additionally satisfies conditional positive homogeneity (CPH).

We define acceptance (rejection) consistency following Weber [136], [137]. The same definition is also known as weak acceptance (rejection) consistency; see [1] among others.

2.14 Definition. A dynamic convex risk measure in discrete time is called *acceptance (rejection) consistent* if for any $X \in \mathcal{X}$ the following condition holds:

$$\rho_t(X) \le (>) 0 \quad \text{if} \quad \rho_{t+1}(X) \le (>) 0 \quad \text{for all } t \in \{t_0 = 0, t_1, ..., t_{N-1}\}.$$

A dynamic convex risk measure in discrete time is called *weakly time consistent* if it is both, acceptance and rejection consistent.

Acceptance consistency formalizes that any financial position which is acceptable at $t+1$ should also be acceptable at the earlier time t. Rejection consistency ensures that at each period of time one stays on the safe side when updating the risk evaluation - a property which seems to be particularly suitable for regulatory use; see [1, page 2]. We refer to [1], Artzner, Delbaen, Eber, Heath and Ku [9], Penner [112], Tutsch [131], [132] and [136], [137] for further studies on acceptance and rejection consistency of dynamic risk measures.

2.15 Definition. A dynamic convex risk measure is called *time consistent* if for any $X,Y \in \mathcal{X}$ the following condition holds:

$$\rho_t(X) \le \rho_t(Y) \quad \text{if} \quad \rho_{t+1}(X) \le \rho_{t+1}(Y) \quad \text{for all } t \in I \text{ with } t \le T-1.$$

This is equivalent to:

$$\text{For all } X \in \mathcal{X}: \ \rho_s(X) = \rho_s(-\rho_t(X)) \quad \text{for all } s,t \in I \text{ with } s \le t \tag{2.2}$$

which, broadly speaking, means that the same risk is assigned to X regardless of whether it is calculated over two time periods or in two steps backwards in time. For dynamic risk measures in discrete time we can give some equivalent characterizations of time consistency:

2.16 Remark. (a) From (2.2) it follows that for time consistent risk measures, for all $t \in I$, we have $\rho_0(X) = \rho_0(-\rho_t(X))$, i.e., in order to quantify the riskiness of X at the initial time 0, it is equivalent to either use the static risk measure ρ_0 directly, i.e., computing $\rho_0(X)$, or to first evaluate the riskiness $\rho_t(X)$ of X at time t and then quantify the risk of $-\rho_t(X)$ at time 0 through $\rho_0(-\rho_t(X))$.

(b) The recursive property (2.2) is crucial for relating time consistent DDRMs to CDRMs in a Brownian setting; see Chapter 7.

2.5 Risk measures via g–expectations

In this section we present the concept of g-expectations and backward stochastic differential equations in the context of risk measures. As pointed out in Peng [111] and Rosazza Gianin [121], among others, it is possible to obtain a large number of static and dynamic risk measures by solving a BSDE with driver g.

For this section we assume $I = [0,T]$, i.e., we only consider the case of CDRMs, and we assume the financial positions to be square-integrable,i.e., $X \in L^2 := L^2(\Omega,\mathcal{F},P)$. Let $(W_t)_{t\in I}$ be a standard one-dimensional Brownian motion on $(\Omega,\mathcal{F},P)$. By $(\mathcal{F}_t^W)_{t\in I}$ we denote the filtration generated by W and by $(\mathcal{F}_t)_{t\in I}$ we denote the augmented filtration associated with $(\mathcal{F}_t^W)_{t\in I}$. For a fixed $T > 0$ let $L^2_{\mathcal{F}}(T;\mathbb{R}^n)$ denote the space of all $\mathbb{R}^n$-valued, adapted processes $(V_t)_{t\in I}$ such that $\mathbb{E}\left[\int_0^T \|V_t\|^2\,dt\right] < \infty$, where $\|\cdot\|$ stands for the Euclidean norm on $\mathbb{R}^n$.

Consider a functional $g : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$, $(\omega,t,y,z) \mapsto \mathbb{R}$. Further on, we write $g(t,y,z)$ instead of $g(\omega,t,y,z)$ and g is supposed to satisfy the following *usual assumptions*:

1. g is Lipschitz in (y,z): There exists a constant $C > 0$ such that for all $t \in I$, for all $(y_0,z_0), (y_1,z_1) \in \mathbb{R} \times \mathbb{R}^d$:

$$|g(t,y_0,z_0) - g(t,y_1,z_1)| \le C\left(|y_0 - y_1| + \|z_0 - z_1\|\right).$$

2. $g(\cdot,y,z) \in L^2_{\mathcal{F}}$ for all $y \in \mathbb{R}$, $z \in \mathbb{R}^d$.

3. For all $t \in I$, $y \in \mathbb{R}:\ g(t,y,0) = 0$.

Under these assumptions on g, for every $X \in L^2$ the BSDE

$$\begin{aligned} -dY_t &= g(t,Y_t,Z_t)dt - Z_t dW_t, \quad \text{for all } t \in I \\ Y_T &= X \end{aligned} \tag{BSDE}$$

has a unique solution $(Y_t,Z_t)_{t\in I} \in L^2_{\mathcal{F}}(T;\mathbb{R}) \times L^2_{\mathcal{F}}(T;\mathbb{R}^d)$. We are particularly interested in the component $(Y_t)_{t\in I}$ of the solution.

2.17 Definition. For any $X \in L^2$, let $(Y_t^X,Z_t^X)_{t\in I} \in L^2_{\mathcal{F}}(T;\mathbb{R}) \times L^2_{\mathcal{F}}(T;\mathbb{R}^d)$ be the solution of (BSDE).

(i) The *g-expectation* $\mathcal{E}_g$ *of* X is defined by

$$\mathcal{E}_g[X] := Y_0^X.$$

(ii) For any $t \in I$ the *conditional g-expectation of* X *under* $\mathcal{F}_t$ is defined by

$$\mathcal{E}_g[X \mid \mathcal{F}_t] := Y_t^X.$$

Let g satisfy the usual assumptions and set $\rho^g : L^2 \to \mathbb{R}$ as follows:

$$\rho^g(X) := \mathcal{E}_g[-X], \quad \text{for all } X \in L^2.$$

We now introduce some additional restrictions on the functional g in order to define when ρ^g is a static risk measure.

2.18 Definition. (i) g is *sublinear* if it is

– *positively homogeneous* in (y,z): For all $t \in I$, for all $\lambda \ge 0$, for all $(y,z) \in \mathbb{R} \times \mathbb{R}^d$:

$$g(t,\lambda y,\lambda z) = \lambda g(t,y,z),$$

– *subadditive* in (y,z): For all $t \in I$, for all $\lambda \ge 0$, for all $(y_0,z_0),(y_1,z_1) \in \mathbb{R} \times \mathbb{R}^d$:

$$g(t,y_0 + y_1,z_0 + z_1) \le g(t,y_0,z_0) + g(t,y_1,z_1).$$

(ii) g is *convex* in (y,z) if for all $t \in I$, for all $(y_0,z_0),(y_1,z_1) \in \mathbb{R} \times \mathbb{R}^d$, for all $\lambda \in (0,1)$:

$$g(t,\lambda y_0 + (1-\lambda)y_1,\lambda z_0 + (1-\lambda)z_1) \le \lambda g(t,y_0,z_0) + (1-\lambda)g(t,y_1,z_1).$$

2.19 Proposition (Proposition 11 from [121])**.** *(i) If g satisfies sublinearity, then ρ^g is a coherent risk measure.*

(ii) If g satisfies convexity, then ρ^g is a convex risk measure.

Let us now set

$$\rho_t^g(X) := \mathcal{E}_g[-X \mid \mathcal{F}_t], \text{ for all } X \in L^2 \text{ and all } t \in I,$$

where g satisfies the usual assumptions. Then the following proposition holds:

2.20 Proposition (Proposition 19 from [121]).

(i) *If g satisfies sublinearity (for instance, $g(z) = \mu \|z\|$ with $\mu > 0$), then $(\rho_t^g)_{t\in I}$ is a coherent and time consistent risk measure.*

(ii) *If g satisfies convexity, then $(\rho_t^g)_{t\in I}$ is a dynamic convex and time consistent risk measure.*

As the functional g is crucial for the construction of risk measures via g-expectations, let us recall some remarks from [121] on how g can be interpreted. g may depend on the dynamics of a stochastic process whose final value is X and g could depend on the preferences of the investor. In particular, the bigger g, the more conservative is the corresponding risk measure and if g is convex (sublinear) then also the corresponding risk measure is convex (coherent). We refer to [121, Section 3] for a more detailed analysis of these relations.

CHAPTER 3

Utility-based risk measures

This chapter gives an overview of some fundamental aspects of the theory of utility functions and possible connections of this theory with risk measures. It may therefore be considered as a preliminary chapter providing necessary results for the main contributions of this thesis in Chapter 4 and Chapter 5.

In the course of this work there are many connections to the classical approach to individuals' attitudes towards risk and risky payoffs via the notion of utility functions, which dates back at least to Bernoulli's work in the 18th century. The theory of risk preferences, their numerical representations, and in particular expected utility was advanced significantly by Arrow [6], Pratt [114] and von Neumann and Morgenstern [134]; see also the monographs of Fishburn [63, 64], Gollier [81], Mas-Colell, Whinston and Green [105], among others. We also refer to [69, Chapter 2] and the references therein for a mathematical perspective on the theory of economic preferences.

In Section 3.1 we review some well-known classes of utility-based risk measures and point out their respective crucial properties. Section 3.2 includes the basic definition of utility functions for our further work and some associated characteristics of utility.

3.1 Already established utility-based risk measures

For $X \in \mathcal{X}$, $u : \mathbb{R} \to \mathbb{R}$ being a strictly increasing, continuous and concave utility function, the certainty equivalent of X is defined as follows:

$$C_u(X) := u^{-1} \left(\mathbb{E} \left[u(\beta X) \right] \right).$$

$\mathbb{C}_u[X]$ is by definition the sure amount that makes an investor indifferent between investing in X and receiving $\mathbb{C}_u[X]$. Certainty equivalents are used to determine the amount of certain wealth which has the same utility as the expected utility of the unknown outcome of X. This is a particularly meaningful approach for random variables X which can attain very large (negative) values (but not $-\infty$) with a small probability (say X can attain -1,

-2, -3, -4 or $-1{,}000{,}000$) such that the expectation of X would not be a meaningful choice for the risk involved in X. Classical risk measures such as value at risk or average value at risk are not perfectly suited either to emphasize the risk of an extreme loss of such a position. A more sensible way of measuring the risk of X is to consider the individual utility of every possible outcome of X and then average these utilities.

The following properties of C_u are immediate consequences of its definition. For any $X \in \mathcal{X}$,

(i) $C_u(m) = \beta m$ for all $m \in \mathbb{R}$.

(ii) $C_u(X)$ is invariant to affine transformations in u.

(iii) $C_u(X) \leq \mathbb{E}[X]$; due to Jensen's inequality.

(iv) $C_u(X) \leq C_u(Y)$ if $X \leq Y$.

Note that $-C_u$ is monotone in the sense of Definition 2.1 and even cash-invariant, i.e., $-C_u(X + m) = -C_u(X) - \beta m$ if u is either linear or of exponential form; see also de Finetti [42], [69, Proposition 2.46] and Nagumo [109]. For $u(x) = a - b\exp(-\gamma x)$, $a \in \mathbb{R}$, $b, \gamma \in \mathbb{R}_+$, $-C_u$ is the entropic risk measure ρ^{ent} from Example 2.6. As noted by Föllmer and Knispel [65], ρ^{ent} is the only risk measure which is, up to a change of sign, at the same time a certainty equivalent with respect to some increasing concave utility function u.

We have just outlined a first connection between risk measures and certainty equivalents (and thereby utility functions). Föllmer and Schied [70] point out an early implicit appearance of risk measures in the theory of preferences: Gilboa and Schmeidler [79] showed that a natural relaxation of the "axioms of rationality" which are formulated in [123] and [134] implies that the linear expected utility under the probability measure P should be replaced by a general coherent risk measure ρ:

$$-\rho(u(X)) = \inf_{Q \in \mathcal{Q}^{\max}} \mathbb{E}_Q[u(X)].$$

In an even further relaxed axiomatic setting which is considered in Maccheroni, Marinacci and Rustichini [100], linear expected utility is generalized by

$$-\rho(u(X)) = \inf_{Q \in \mathcal{M}_1(P)} (\mathbb{E}_Q[u(X)] + \alpha^{\min}(Q)),$$

thus ρ is now a convex risk measure; see Theorem 2.7. Here, concavity of u can be interpreted as aversion of model uncertainty; see Föllmer, Schied and Weber [71] and [100].

However, a basic idea of this thesis is to make use of "utility" functions in the sense of [134] in the context of risk measurement when the probability measure is a priori known. In the following we review some existing notions developed in this spirit.

3.1 Example. If $u : \mathbb{R} \to \mathbb{R}$ is an increasing, concave function that is not identically constant, the *utility-based shortfall risk* is introduced in [67] as

$$\rho^{\text{SR}}(X) := \inf \{\eta \in \mathbb{R} : \mathbb{E}[u(\beta X + \eta)] \geq \lambda\},$$

where λ is a fixed value in the interior of the range of u. ρ^{SR} fulfills (M), (CI), (N) and (C) from Definition 2.1, i.e., it is a convex risk measure. However, ρ^{SR} fulfills (PH) from

Definition 2.1 only if u is piecewise linear with a kink at 0; see Weber [137, Corollary 3.2]. As shown in [137], $\rho^{\mathrm{SR}}(X)$ are the only law-invariant convex risk measures which are also invariant under randomization. In Chapter 6 we get more into detail on this property.

3.2 Example. For a strictly increasing, continuous function $u : \mathbb{R} \to \mathbb{R}$ the *certainty equivalent as risk measure* is defined in [69, Example 4.13] via

$$\rho^{\mathrm{CE}}(X) := -u^{-1}\left(\mathbb{E}\left[u\left(\beta X\right)\right]\right).$$

ρ^{CE} is monotone and normalized, but ρ^{CE} is cash invariant and convex only if u is either linear or exponential, the latter leading to the entropic risk measure. $\rho^{\mathrm{CE}}(X)$ preserves convex stochastic ordering; see [108, Theorem 2.8].

3.3 Example. For a concave, increasing function $u : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ with $u(0) = 0$ and $u(x) \leq x$ for all $x \in \mathbb{R}$, the *optimized certainty equivalent* is introduced in [18, 19] as

$$C^{\mathrm{OCE}}(X) := \sup_{\eta \in \mathbb{R}} \{-\eta + \mathbb{E}[u(\beta X + \eta)]\}.$$

Its negative counterpart $\rho^{\mathrm{OCE}}(X) := -C^{\mathrm{OCE}}(X)$ is a convex risk measure. ρ^{OCE} is not coherent unless u is piecewise linear with a kink at 0.

3.4 Example. For a random variable X that represents costs or losses, Krokhmal [97] defines the functional

$$\rho(X) := \inf_{\eta \in \mathbb{R}} \{\eta + \phi(\beta X - \eta)\},$$

which is a lower semi-continuous coherent risk measure provided the abstract functional $\phi : \mathcal{X} \to \mathbb{R} \cup \{\infty\}$ satisfies monotonicity, convexity, positive homogeneity, lower semi-continuity and is such that $\phi(\eta) > \eta$ for all $\eta \in \mathbb{R}$, $\eta \neq 0$. Vinel and Krokhmal [133] specify ϕ in terms of an increasing, convex disutility function $v : \mathbb{R} \to \mathbb{R}$ and a penalty parameter $\alpha \in (0,1)$ via

$$\phi(X) = \tfrac{1}{1-\alpha} v^{-1}\left(\mathbb{E}\left[v(X)\right]\right).$$

This leads to *certainty equivalent measures of risk*

$$\rho^{\mathrm{CEM}}(X) = \min_{\eta \in \mathbb{R}} \left\{\eta + \tfrac{1}{1-\alpha} v^{-1}\left(\mathbb{E}\left[v\left(\beta X - \eta\right)\right]\right)\right\}.$$

3.2 Utility functions

In the course of this thesis we focus on utility functions in the sense of the following definition:

3.5 Definition. A *utility function* is a C^3 function $u : \mathbb{R}_+ \to \mathbb{R}$ such that $u'(x) > 0$ and $u''(x) < 0$ for all $x \in \mathbb{R}_+$, and such that the absolute risk tolerance τ_u is concave. Here $\tau_u : \mathbb{R}_+ \to \mathbb{R}$ is given by

$$\tau_u(x) := -\frac{u'(x)}{u''(x)}, \quad x \in \mathbb{R}_+.$$

We denote by $\mathcal{U}$ the set of all utility functions. For $u \in \mathcal{U}$ we set $u(0) := \lim_{x \downarrow 0} u(x) \in [-\infty,\infty)$, $u(\infty) := \lim_{x \uparrow \infty} u(x) \in (-\infty,\infty]$ and regard u as a mapping $u : [0,\infty] \to [-\infty,\infty]$ without further notice.

As pointed out in Gollier [81, page 114], concave absolute risk tolerance is a natural assumption for utility functions. For instance, in a classical investment context, it means that the reduction of risk premia due to an increase in wealth is a decreasing function of wealth; on the other hand, absolute risk aversion cannot be positive, decreasing and concave everywhere. In fact, most of the commonly used utility functions, including power, logarithmic and exponential utilities, exhibit concave absolute risk tolerance. Moreover, Ben-Tal and Teboulle [19, Corollary 5.1] show that τ_u is concave if and only if the associated certainty equivalent functional $X \mapsto u^{-1}(\mathbb{E}[u(X)])$ is concave.

3.6 Definition. For $u \in \mathcal{U}$ we define the left inverse $u^{-1} : \mathbb{R} \to [0,\infty] \cup \{-\infty\}$ via

$$u^{-1}(t) := \inf\{x \in \mathbb{R}_+ : u(x) \geq t\}, \quad u(0) \leq t < u(\infty),$$

and $u^{-1}(t) := -\infty$ for $t < u(0)$, $u^{-1}(t) := \infty$ for $t \geq u(\infty)$. Then u^{-1} is right-continuous and $u^{-1}(u(x)) = x$ for all $x \in [0,\infty]$.

3.7 Remark. u being strictly increasing reflects the common sense assumption that investors always prefer more to less. The investor is assumed to be risk averse which is captured by concavity of u. Risk aversion implies that for any random variable X, its expectation $\mathbb{E}[X]$ is preferred to X,

$$u(\mathbb{E}[X]) \geq \mathbb{E}[u(X)],$$

which in turn is just Jensen's inequality for u, i.e., is equivalent to say that u is concave. The positive difference $\mathbb{E}[X] - C_u(X)$ is called the investor's *risk premium* which may be interpreted as the amount that an individual is willing to pay in order to avoid a risk. $u'(x)$ measures the marginal improvement in utility with changes in x. $u''(x)$ is the rate of change of the satisfaction: It gets harder and harder to increase satisfaction as x increases.

In general, for any utility function $u \in \mathcal{U}$ the *Arrow-Pratt coefficient of absolute risk aversion (ARA)* of u at level x is defined as

$$\varrho_u(x) = -\frac{u''(x)}{u'(x)};$$

see also [6] and [114]. We denote by $\mathcal{U}_{\mathrm{IARA}}$ the class of utility functions $u \in \mathcal{U}$ that feature increasing absolute risk aversion, by $\mathcal{U}_{\mathrm{CARA}}$ those with constant absolute risk aversion, and by $\mathcal{U}_{\mathrm{DARA}}$ those with decreasing absolute risk aversion. The *relative risk aversion (RRA)* is defined as

$$x\varrho_u(x).$$

We denote by $\mathcal{U}_{\mathrm{IRRA}}$ the set of utility functions $u \in \mathcal{U}$ with increasing relative risk aversion, and by $\mathcal{U}_{\mathrm{CRRA}}$ those with constant relative risk aversion.

Standard examples of classes of utility functions are all $u \in \mathcal{U}$ with

(i) constant absolute risk aversion (CARA): $\varrho_u(x) = \gamma$ for all $x \in D_u$.

(ii) constant relative risk aversion (CRRA): $x\varrho_u(x) = \gamma$ for all $x \in D_u$.

Any CARA utility function $u(x)$ implies the same behavior for all $x \in \mathbb{R}$ thus, for example, in a standard portfolio problem with one risk-free asset and one risky asset the optimal holding of the risky asset is independent of the investor's initial wealth, i.e., it does not increase even if the investor experiences an increase in wealth. Thus CARA exponential utility which is convenient for calculation is considered a rarely plausible choice in reality; see [81] and [114].

The relative risk aversion on the other hand is a well-known characteristic of individuals' risk preferences. Recent studies on the risk attitudes of representative investors such as the works by Chiappori and Paiella [37], Friend and Blume [73], [107] or Szpiro [129] make a case for constant (or slightly increasing) relative risk aversion. CRRA is implied by decreasing absolute prudence, i.e., $-u'''(x)/u''(x)$ decreasing in x, which is widely considered a reasonable condition; see, e.g., Kimball [92], [93]. In the standard portfolio problem with one risk-free asset and one risky asset, if the portfolio manager has constant relative risk aversion, he will choose to keep the fraction of the portfolio held in the risky asset unchanged if he experiences an increase in wealth. Note that all CRRA utility functions exhibit DARA, since

$$\varrho_u'(x) = -\frac{\gamma}{x^2},$$

where $\gamma \in \mathbb{R}_+$ denotes the investor's relative risk aversion. We refer to Chapter 4 and Chapter 5, where utility functions with constant relative risk aversion are of particular importance. One can check that, due to the fact that any utility function $u \in \mathcal{U}$ has concave absolute risk tolerance, it is impossible for $u \in \mathcal{U}$ to have decreasing relative risk aversion. These are the cases that can actually occur:

1. If $\tau_u(x)$ is concave and decreasing, then $u \in \mathcal{U}_{\text{IRRA}}$.

2. If $\tau_u(x)$ is concave and constantly positive, then $u \in \mathcal{U}_{\text{IRRA}}$.

3. $\tau_u(x)$ is concave and increasing and $u \in \mathcal{U}_{\text{IRRA}}$.

Let us now argue why it is impossible that (a) $\tau_u(x)$ is concave and constantly negative (then $u \in \mathcal{U}_{\text{DRRA}}$), or that (b) $\tau_u(x)$ is concave and increasing and $u \in \mathcal{U}_{\text{DRRA}}$, respectively. For $u \in \mathcal{U}$ we have $u'(x) > 0$ and $u''(x) < 0$ for all $x \in \mathbb{R}_+$. Thus $\tau_u(x) > 0$ holds for any $x \in \mathbb{R}_+$ and (a) cannot occur. Now, let us assume that $\tau_u(x)$ is concave and increasing. Then $\tau_u' \geq 0$ and τ_u' is decreasing, and due to the mean value theorem, we know that for any $x > 0$ there exists $\xi \in (0,x)$ such that

$$\tau_u'(\xi) = \frac{\tau_u(x) - \tau_u(0+)}{x}.$$

Since τ_u' is decreasing and τ_u is positive for any $x > 0$:

$$\tau_u'(x) \leq \frac{\tau_u(x) - \tau_u(0+)}{x} \leq \frac{\tau_u(x)}{x}$$

which is equivalent to

$$\tau_u(x) \geq x\tau_u'(x).$$

The above implies that if $\tau_u(x)$ is concave and increasing, then:

$$(x\varrho_u(x))' = \left(\frac{x}{\tau_u(x)}\right)' = \frac{\tau_u(x) - x\tau_u'(x)}{\tau_u(x)^2} \geq 0 \quad \text{for all } x > 0.$$

Therefore τ_u being concave and increasing always leads to $u \in \mathcal{U}_{\text{IRRA}}$ (case 3.) and (b) cannot occur.

We conclude this section with a result that allows us to relate the "degree of concavity" of utility functions to their respective coefficient of absolute risk aversion and their respective risk premiums:

3.8 Proposition (Proposition 2.44 from [69]). *Let $u,v \in \mathcal{U}$. The following conditions are equivalent:*

(i) $\varrho_u(x) \geq \varrho_v(x)$ *for all* $x \in \mathbb{R}_+$.

(ii) $u = g \circ v$ *for a strictly increasing concave function* g.

(iii) $C_v(X) \geq C_u(X)$ *for all* $X \in \mathcal{X}$.

We note that a more concave utility function relates to a greater absolute risk aversion and to a higher risk premium which is intuitively understandable; see also [114, Theorem 1].

CHAPTER 4

Optimal expected utility risk measures

In this chapter, which is an extended version of Geissel, Sass and Seifried [77] we propose a construction of risk measures based on certainty equivalents which we call *optimal expected utility (OEU) risk measures*. In a nutshell, OEU is based on an asset allocation problem where the agent decides endogenously how much capital η to set aside as a reserve today if he is to receive a given stochastic payoff X tomorrow. The first-order condition for the underlying optimization reads

$$\alpha \mathbb{E}\left[u'\left(X+\eta^*\right)\right] = \beta u'(x),$$

where $x := u^{-1}(\mathbb{E}[u(X+\eta^*)])$. Thus the (expected) marginal utility of the risky payoff including reserves, $X+\eta^*$, equals the marginal utility of today's certainty equivalent of that same quantity.

This chapter is organized as follows: In Section 4.1 we introduce OEU risk measures and investigate the link between OEU and the underlying utility function. Sections 4.2, 4.3 and 4.4 include various properties of OEU; in particular, Theorem 4.3 establishes OEU as a convex risk measure, and Theorem 4.16 gives conditions for when it is coherent. In Section 4.5 we develop a method for recovering the utility function from a given OEU and in Section 4.6 we give a dual representation of OEU. Finally, Section 4.7 presents applications of OEU to some simplified problems in risk management.

Our notion of optimal expected utility risk measures as developed below is inspired by by the previous studies on utility-based risk measures which we recalled in Chapter 3. Its distinctive feature is that it takes seriously the idea that the "utility" function u captures the investor's utility in the sense of classical utility theory. In particular, OEU is able to deal with standard utility functions, including power and exponential utilities, and we are able to link properties of OEU to corresponding properties of the underlying utility function. In addition, we provide a decision-theoretic foundation of OEU in terms of a well-defined certainty equivalent optimization problem. In particular, unlike ρ^{OCE}, OEU strictly distinguishes between "cash" ("dollar") and "utility" quantities.

4.1 Introducing optimal expected utility risk measures

In this section we introduce optimal expected utility risk measures. We further give specific examples of OEU for some commonly used utility functions, develop the main properties of OEU, and study its relationship to other risk measures.

4.1 Definition. The *optimal expected utility* (OEU) risk measure is defined as the map $\rho^u : \mathcal{X} \to \mathbb{R}$,

$$\rho^u(X) := -\sup_{\eta \in \mathbb{R}} \left\{ -\beta\eta + \alpha u^{-1} \left(\mathbb{E}\left[u\left(X + \eta \right) \right] \right) \right\}, \tag{OEU}$$

where $\alpha > 0$ captures the investor's subjective time preference and by convention $\mathbb{E}[u(Y)] := -\infty$ if $P(Y < 0) > 0$.

To understand the economic rationale underlying OEU, consider an investor that is able to borrow and invest at the market interest rate, and aims to maximize the sum of capital available today and the certainty equivalent of his capital in the future (the latter discounted by his subjective rate of time preference). If the investor holds a financial position with net value X and decides on the amount η that he is to borrow to allocate capital optimally over time, he formally faces the problem to maximize

$$H : [-X_{\min}, \infty) \to [-\infty, \infty), \quad H(\eta) := -\beta\eta + \alpha u^{-1} \left(\mathbb{E}\left[u\left(X + \eta \right) \right] \right) \tag{4.1}$$

over η, i.e., the quantity inside the supremum in (OEU).[1] The risk measure $\rho^u(X)$ then represents the amount of capital required to make investment in X appear favorable to him (equivalently, $-\rho^u(X)$ is the maximal amount he would be willing to pay in order to obtain X). Note that at the optimum the investor always puts aside a sufficient amount to cover all future losses in X, thus guaranteeing to avoid bankruptcy. Of course, this construction is inspired by a general correspondence between certainty equivalents, risk measures, and indifference prices; see, e.g., Barrieu and El Karoui [13, page 81]: For an arbitrary risk measure ρ, the quantity $\rho(X)$ can be seen "as the opposite of the 'buyer's indifference price' of this position, since when paying the amount $-\rho(X)$, the new exposure $X - (-\rho(X))$ does not carry any risk with positive measure, i.e., the agent is somehow indifferent using this criterion between doing nothing and having this 'hedged' exposure. The convex risk measures appear therefore as a natural extension of utility functions as they can be seen directly as an indifference pricing rule."

4.2 Remark. To rule out unbounded leveraging, we assume that

$$\alpha < \beta,$$

unless explicitly stated otherwise. This means that the investor's rate of time preference exceeds the risk-free interest rate in the financial market. If this condition is violated, the investor's preferences degenerate in the sense that he will, already in the absence of

[1] For $P(X + \eta < 0) > 0$ we have $\mathbb{E}[u(X + \eta)] = -\infty$, hence $H(\eta) = -\infty$ is not the optimum and it suffices to consider $\eta \in [-X_{\min}, \infty)$.

risk, allocate as much capital as possible to future consumption and thus typically aim to borrow unlimited amounts: Indeed, for $\alpha > \beta$ we have for any deterministic payoff $m \in \mathbb{R}$

$$\rho^u(m) = -\sup_{\eta\in\mathbb{R}}\{-\beta\eta + \alpha(m+\eta)\} = -\infty.$$

The limiting case $\alpha = \beta$ is investigated below; see Proposition 4.9, Proposition 4.14 and Remark 4.23.

4.2 Main properties and examples of OEU

The following theorem demonstrates that OEU satisfies all the axioms of a (convex) risk measure.

4.3 Theorem. *For any $u \in \mathcal{U}$, ρ^u is a convex risk measure.*

Proof. (M) is an obvious conclusion of the definition of OEU and the fact that u and u^{-1} are both increasing functions: Let $X,Y \in \mathcal{X}$ such that $X \le Y$, then also $X_{\min} \le Y_{\min}$, thus

$$\begin{aligned}
\rho^u(X) &= -\sup_{\eta>-X_{\min}} \left\{-\beta\eta + \alpha u^{-1}\left(\mathbb{E}\left[u\left(X+\eta\right)\right]\right)\right\} \\
&\ge -\sup_{\eta>-X_{\min}} \left\{-\beta\eta + \alpha u^{-1}\left(\mathbb{E}\left[u\left(Y+\eta\right)\right]\right)\right\} \\
&\ge -\sup_{\eta>-Y_{\min}} \left\{-\beta\eta + \alpha u^{-1}\left(\mathbb{E}\left[u\left(Y+\eta\right)\right]\right)\right\} \\
&= \rho^u(Y).
\end{aligned}$$

Cash invariance follows directly from the definition of OEU: Let $m \in \mathbb{R}$, then

$$\begin{aligned}
\rho^u(X+m) &= -\sup_{\eta>-X_{\min}-m} \left\{-\beta\eta + \alpha u^{-1}\left(\mathbb{E}\left[u\left(X+m+\eta\right)\right]\right)\right\} \\
&= -\left(\beta m + \sup_{\eta>-X_{\min}-m} \left\{-\beta(\eta+m) + \alpha u^{-1}\left(\mathbb{E}\left[u\left(X+(\eta+m)\right)\right]\right)\right\}\right) \\
&\overset{\eta':=(\eta+m)}{=} -\left(\beta m + \sup_{\eta'>-X_{\min}} \left\{-\beta\eta' + \alpha u^{-1}\left(\mathbb{E}\left[u\left(X+\eta'\right)\right]\right)\right\}\right) \\
&= -\beta m + \rho^u(X).
\end{aligned}$$

To show (N) note that for all $0 < \alpha < \beta$:

$$\begin{aligned}
\rho^u(0) &= -\sup_{\eta\in\mathbb{R}} \left\{-\beta\eta + \alpha u^{-1}\left(\mathbb{E}\left[u\left(0+\eta\right)\right]\right)\right\} \\
&= -\sup_{\eta>0} \left\{\underbrace{(\alpha-\beta)}_{<0}\eta\right\} \\
&= 0.
\end{aligned}$$

To establish convexity, let $\lambda \in [0,1]$ and define

$$g(x_0,x_1) := u\left(\lambda u^{-1}(x_0) + (1-\lambda)u^{-1}(x_1)\right), \quad (x_0,x_1) \in \mathbb{R} \times \mathbb{R}.$$

Further denote:

$$u' := u'(\lambda u^{-1}(x_0) + (1-\lambda)u^{-1}(x_1)), \quad u_0' := u'(u^{-1}(x_0)), \quad u_1' := u'(u^{-1}(x_1)),$$
$$u'' := u''(\lambda u^{-1}(x_0) + (1-\lambda)u^{-1}(x_1)), \quad u_0'' := u''(u^{-1}(x_0)), \quad u_1'' := u''(u^{-1}(x_1)).$$

Concavity of τ_u yields

$$\underbrace{\tau_u(\lambda u^{-1}(x_0) + (1-\lambda)u^{-1}(x_1))}_{=-\frac{u'}{u''}} \geq \lambda \underbrace{\tau_u(u^{-1}(x_0))}_{=-\frac{u_0'}{u_0''}} + (1-\lambda) \underbrace{\tau_u(u^{-1}(x_1))}_{=-\frac{u_1'}{u_1''}}.$$

By multiplying the above equation by the negative factor $-\lambda u'' u_0'' \frac{1}{(u_0')^3}$, we get

$$-\lambda u' u_0'' \frac{1}{(u_0')^3} \geq -\lambda^2 \frac{1}{(u_0')^2} u'' + \underbrace{(1-\lambda)\frac{u_1'}{u_1''}(-\lambda)u'' u_0'' \frac{1}{(u_0')^3}}_{\geq 0},$$

thus

$$-\lambda u' u_0'' \frac{1}{(u_0')^3} \geq -\lambda^2 \frac{1}{(u_0')^2} u''$$

which implies

$$\underbrace{u'' \frac{\lambda^2}{(u_0')^2} - u' \frac{\lambda u_0''}{(u_0')^3}}_{=\frac{\partial^2 g}{\partial x_0^2}} \geq 0.$$

Further $q := \lambda(1-\lambda)u_0'' u_1'' u' u'' (u_0' u_1')^{-3} \leq 0$ and therefore, due to concavity of τ_u we have

$$\tau_u(\lambda u^{-1}(x_0) + (1-\lambda)u^{-1}(x_1)) \geq \lambda \tau_u(u^{-1}(x_0)) + (1-\lambda)\tau_u(u^{-1}(x_1)),$$

thus, by definition of τ_u,

$$q\frac{u'}{u''} \geq q(\lambda \frac{u_0'}{u_0''} + (1-\lambda)\frac{u_1'}{u_1''}).$$

Inserting q yields

$$\frac{\lambda(1-\lambda)u_0'' u_1'' (u')^2}{(u_0' u_1')^3} \geq \frac{\lambda^2(1-\lambda)u_1'' u' u''}{(u_0')^2 (u_1')^3} + \frac{\lambda(1-\lambda)^2 u_0'' u' u''}{(u_0')^3 (u_1')^2},$$

and, therefore

$$u'' \left(\frac{\lambda}{u_0'}\right)^2 \left(-u' \frac{(1-\lambda)u_1''}{(u_1')^3}\right) - u' \frac{\lambda u_0''}{(u_0')^3} u'' \left(\frac{1-\lambda}{u_1'}\right)^2 + u' \frac{\lambda u_0''}{(u_0')^3} u' \frac{(1-\lambda)u_1''}{(u_1')^3} \geq 0.$$

Hence the determinant of the Hessian matrix H of g is positive:

$$\det(H) \geq 0.$$

This implies that g is convex, and Jensen's inequality implies that for $Y_0, Y_1 \in \mathcal{X}$

$$\begin{aligned}
\mathbb{E}[u(\lambda Y_0 + (1-\lambda)Y_1)] &= \mathbb{E}[g(u(Y_0), u(Y_1))] \\
&\geq g\left(\mathbb{E}[u(Y_0)], \mathbb{E}[u(Y_1)]\right) \\
&= u\left(\lambda u^{-1}(\mathbb{E}[u(Y_0)]) + (1-\lambda)u^{-1}(\mathbb{E}[u(Y_1)])\right).
\end{aligned}$$

It follows that

$$\begin{aligned}
\rho^u(\lambda X_0 + (1-\lambda)X_1) &= -\sup_{\eta\in\mathbb{R}} \left\{-\beta\eta + \alpha u^{-1}\left(\mathbb{E}\left[u\left(\lambda X_0 + (1-\lambda)X_1 + \eta\right)\right]\right)\right\} \\
&= -\sup_{\eta_0,\eta_1\in\mathbb{R}} \Big\{ -\lambda\beta\eta_0 - (1-\lambda)\beta\eta_1 \\
&\qquad + \alpha u^{-1}\left(\mathbb{E}\left[u\left(\lambda(X_0+\eta_0) + (1-\lambda)(X_1+\eta_1)\right)\right]\right)\Big\} \\
&\leq -\sup_{\eta_0,\eta_1\in\mathbb{R}} \Big\{ -\lambda\beta\eta_0 + \alpha\lambda u^{-1}\left(\mathbb{E}[u(X_0+\eta_0)]\right) \\
&\qquad - (1-\lambda)\beta\eta_1 + \alpha(1-\lambda)u^{-1}\left(\mathbb{E}[u(X_1+\eta_1)]\right)\Big\} \\
&= \lambda\rho^u(X_0) + (1-\lambda)\rho^u(X_1).
\end{aligned}$$

This implies that ρ^u is convex, and the proof is complete. □

4.4 Remark. Theorem 4.3 implies that ρ^u satisfies (CI) and (N) and hence in particular preserves cash:

$$\rho^u(m) = -\beta m, \quad m \in \mathbb{R}.$$

(CI) is an indispensable property of risk measures, but it is this condition that is in general not satisfied by traditional utility-based risk measures such as $\rho^{\mathrm{CE}}(X) = -u^{-1}(\mathbb{E}[u(\beta X)])$. In the context of OEU, (CI) is ensured by the special construction of (OEU) via balancing the present value $\beta\eta$ of the amount borrowed against the discounted certainty equivalent of the risky payoff, $\alpha u^{-1}\left(\mathbb{E}\left[u(\cdot)\right]\right)$. A related concept in an abstract vector space setting is the *infimal convolution* of two functions $f,g : \mathcal{H} \to \mathbb{R}$ defined on a vector space $\mathcal{H}$ via

$$f \,\square\, g : \mathcal{H} \to \mathbb{R} : \; x \mapsto \inf_{y\in\mathcal{X}} \left\{f(y) + g(x-y)\right\};$$

see, e.g., Bauschke and Combettes [16, Chapter 12]. The function $f \,\square\, g$ inherits monotonicity and convexity from f and g, and is always cash invariant.

In the following we investigate OEU in more detail. In particular, we discuss whether and where the supremum in (OEU) is attained, show uniqueness, and analyze the associated first-order condition. For this analysis, we first investigate the function H in (4.1) more closely.

4.5 Lemma. *For every $u \in \mathcal{U}$ and every $X \in \mathcal{X}$, the function H defined in* (4.1) *is concave and of class C^1 on $(-X_{\min}, \infty)$. Moreover, H is continuous at $\eta = -X_{\min}$.*

Proof. Concavity on $(-X_{\min},\infty)$ is an immediate consequence of the proof of Theorem 4.3: We know that

$$g(x_0,x_1) := u\left(\lambda u^{-1}(x_0) + (1-\lambda)u^{-1}(x_1)\right)$$

is convex if τ_u is concave. Therefore, for any $0 \le \lambda \le 1$, $\eta_1,\eta_2 \in \mathbb{R}$:

$$\begin{aligned}
H(\lambda\eta_1 + (1-\lambda)\eta_2) &= -\beta(\lambda\eta_1 + (1-\lambda)\eta_2) + \alpha u^{-1}\left(\mathbb{E}\left[u\left(X + \lambda\eta_1 + (1-\lambda)\eta_2\right)\right]\right) \\
&= -\beta(\lambda\eta_1 + (1-\lambda)\eta_2) + \alpha u^{-1}\left(\mathbb{E}\left[u\left(\lambda(X+\eta_1) + (1-\lambda)(X+\eta_2)\right)\right]\right) \\
&\ge -\lambda\beta\eta_1 + \lambda\alpha u^{-1}\left(\mathbb{E}\left[u(X+\eta_1)\right]\right) \\
&\qquad - (1-\lambda)\beta\eta_2 + (1-\lambda)\alpha u^{-1}\left(\mathbb{E}\left[u(X+\eta_2)\right]\right) \\
&= \lambda H(\eta_1) + (1-\lambda)H(\eta_2).
\end{aligned}$$

Differentiability on $(-X_{\min},\infty)$ follows from the chain rule and dominated convergence, using the fact that, for any fixed $\varepsilon > 0$, $u'(X+\eta)$ is uniformly bounded from above and from below and $\mathbb{E}[u(X+\eta)] \in (u(-X_{\min}+\varepsilon),u(\infty))$ for $\eta \ge -X_{\min}+\varepsilon$. To establish continuity at $\eta = -X_{\min}$, it suffices to show that the function

$$f : [-X_{\min},\infty) \to [-\infty,\infty), \quad f(\eta) := u^{-1}\left(\mathbb{E}\left[u\left(X+\eta\right)\right]\right) \tag{4.2}$$

is continuous at $\eta = -X_{\min}$. Note that $\eta \downarrow -X_{\min}$ if and only if $X+\eta \downarrow X - X_{\min}$; in this case, as $u' > 0$, we have $u(X+\eta) \downarrow u(X - X_{\min})$. Monotone convergence implies

$$\mathbb{E}[u^-(X+\eta)] \uparrow \mathbb{E}[u^-(X - X_{\min})] \in [0,\infty],$$

and dominated convergence yields

$$\mathbb{E}[u^+(X+\eta)] \downarrow \mathbb{E}[u^+(X - X_{\min})] \in [0,\infty).$$

Thus we conclude that

$$\mathbb{E}[u(X+\eta)] \downarrow \mathbb{E}[u(X - X_{\min})] \in [-\infty,\infty).$$

Since $\mathbb{E}[u(X+\eta)] \in (u(0),u(\infty))$ for any $\eta > -X_{\min}$, right-continuity of u^{-1} yields continuity at $\eta = -X_{\min}$. □

Since H is concave by Lemma 4.5, we can formally define

$$H(\infty) := \lim_{\eta\uparrow\infty} H(\eta) \in [-\infty,\infty).$$

We adopt this convention in all that follows to simplify notation. With this convention, Lemma 4.5 implies that the supremum in the definition (OEU) of ρ^u is attained in $[-X_{\min},\infty]$.

To establish uniqueness, we show that H is strictly concave on an interval that contains the maximizer.

4.6 Lemma. *Suppose that $u \in \mathcal{U}$ and $X \in \mathcal{X}$, let the function H be defined as in* (4.1)*, and set*

$$b := \sup\left\{\eta > -X_{\min} : H'(\eta) > \alpha - \beta\right\}.$$

Then H is strictly concave on $[-X_{\min},b]$. In particular, H is strictly concave on $\{H' \ge 0\}$.

Proof. By Lemma 4.5 the function H is concave and of class C^1 on $(-X_{\min},\infty)$. Hence it suffices to show that $H''(\eta) < 0$ whenever $\eta \in (-X_{\min},b)$, i.e., whenever $\eta > -X_{\min}$ is such that

$$H'(\eta) = -\beta + \alpha \frac{\mathbb{E}[u'(X+\eta)]}{u'(u^{-1}(\mathbb{E}[u(X+\eta)]))} > \alpha - \beta.$$

Thus in the following we demonstrate that $H''(\eta) < 0$ for all $\eta > -X_{\min}$ with

$$\mathbb{E}[u'(X+\eta)] > u'(u^{-1}(\mathbb{E}[u(X+\eta)])). \tag{4.3}$$

Since u exhibits concave absolute risk aversion, the mapping $X \mapsto u^{-1}(\mathbb{E}[u(X)])$ is concave. Therefore, with f defined as in (4.2) above and for arbitrary $\lambda \in [0,1]$, we have

$$\begin{aligned} f(\lambda\eta_1 + (1-\lambda)\eta_2) &= u^{-1}\left(\mathbb{E}\left[u(X + \lambda\eta_1 + (1-\lambda)\eta_2)\right]\right) \\ &= u^{-1}\left(\mathbb{E}\left[u(\lambda(X+\eta_1) + (1-\lambda)(X+\eta_2)\right]\right) \\ &\geq \lambda u^{-1}\left(\mathbb{E}\left[u(X+\eta_1)\right]\right) + (1-\lambda)u^{-1}\left(\mathbb{E}\left[u(X+\eta_2)\right]\right) \\ &= \lambda f(\eta_1) + (1-\lambda)f(\eta_2). \end{aligned}$$

Hence f is concave. On the other hand, it follows as in the proof of Lemma 4.5 by dominated convergence that f is twice continuously differentiable on $(-X_{\min},\infty)$. Thus it follows that

$$f''(\eta) = \frac{\mathbb{E}[u''(X+\eta)]u'(u^{-1}(\mathbb{E}[u(X+\eta)])) - \mathbb{E}[u'(X+\eta)]^2 \frac{u''(u^{-1}(\mathbb{E}[u(X+\eta)]))}{u'(u^{-1}(\mathbb{E}[u(X+\eta)]))}}{u'(u^{-1}(\mathbb{E}[u(X+\eta)]))^2} \leq 0$$

or equivalently

$$\mathbb{E}[u''(X+\eta)]u'(u^{-1}(\mathbb{E}[u(X+\eta)])) \leq \mathbb{E}[u'(X+\eta)]^2 \frac{u''(u^{-1}(\mathbb{E}[u(X+\eta)]))}{u'(u^{-1}(\mathbb{E}[u(X+\eta)]))}, \quad \text{for all } \eta > -X_{\min}. \tag{4.4}$$

We conclude that for all $\eta \in (-X_{\min},b)$ (again using dominated convergence to justify the interchange of differentiation and expectation)

$$\begin{aligned} H''(\eta) &= \alpha \frac{\mathbb{E}[u''(X+\eta)]u'(u^{-1}(\mathbb{E}[u(X+\eta)])) - \mathbb{E}[u'(X+\eta)]^2 \frac{u''(u^{-1}(\mathbb{E}[u(X+\eta)]))}{u'(u^{-1}(\mathbb{E}[u(X+\eta)]))}}{u'(u^{-1}(\mathbb{E}[u(X+\eta)]))^2} \\ &< \alpha \frac{\mathbb{E}[u''(X+\eta)]\mathbb{E}[u'(X+\eta)] - \mathbb{E}[u'(X+\eta)]^2 \frac{u''(u^{-1}(\mathbb{E}[u(X+\eta)]))}{u'(u^{-1}(\mathbb{E}[u(X+\eta)]))}}{u'(u^{-1}(\mathbb{E}[u(X+\eta)]))^2} \\ &\leq \alpha \frac{\mathbb{E}[u''(X+\eta)]}{u'(u^{-1}(\mathbb{E}[u(X+\eta)]))^2} \left(\mathbb{E}[u'(X+\eta)] - u'\left(u^{-1}(\mathbb{E}[u(X+\eta)])\right)\right) \\ &< 0, \end{aligned}$$

where the strict inequalities are due to (4.3) and we use (4.4) in the third line. This completes the proof. □

Using Lemmas 4.5 and 4.6, we are in a position to give a precise characterization of the maximizer η^* in the definition (OEU). The first-order condition for an interior optimizer η^* in (OEU) is given by

$$\alpha\mathbb{E}\left[u'(X+\eta^*)\right] = \beta u'(x), \quad \text{where } x := u^{-1}\left(\mathbb{E}\left[u\left(X+\eta^*\right)\right]\right). \tag{4.5}$$

Thus the (expected) marginal utility of $X+\eta^*$ equals the marginal utility of today's certainty equivalent of $X+\eta^*$. The following result provides a precise mathematical formulation.

4.7 Proposition. *Suppose that $u \in \mathcal{U}$ and $X \in \mathcal{X}$, and let $H(\infty) := \lim_{\eta\uparrow\infty} H(\eta)$.*

(i) The supremum in (OEU) *is uniquely attained at $\eta^* \in [-X_{\min},\infty]$ where*

$$\eta^* := \sup\left\{\eta \in (-X_{\min},\infty) \,:\, \alpha\mathbb{E}\left[u'\left(X+\eta\right)\right] \geq \beta u'(x) \text{ with } x := u^{-1}\left(\mathbb{E}\left[u\left(X+\eta\right)\right]\right)\right\}.$$

If $\eta^* \in (-X_{\min},\infty)$, then η^* is the unique solution of the first-order condition (4.5).

(ii) If u satisfies

$$\limsup_{\eta\uparrow\infty} \frac{u'(\eta+a)}{u'(\eta+b)} \leq 1 \quad \text{for all } a,b \in \mathbb{R},$$

then $\eta^* \in [-X_{\min},\infty)$.

(iii) If $u(0) = -\infty$ and $P(X = X_{\min}) > 0$, then $\eta^* \in (X_{\min},\infty]$.

In particular, if the conditions in (ii) and (iii) hold, then the maximizer η^* in (OEU) *is the unique solution of the first-order condition* (4.5).

Proof. (i) If H is strictly increasing on $(-X_{\min},\infty)$, then $b=\infty$ and the supremum of H is attained at $\eta^*=\infty$. Otherwise, if we define b as in Lemma 4.6, then by continuity and concavity, see Lemma 4.5, H attains its maximum in $[-X_{\min},b)$, and strict concavity implies uniqueness of the maximizer.

(ii) Jensen's inequality implies that for $\eta > -X_{\min}$

$$\begin{aligned} H'(\eta) &= -\beta + \alpha\frac{\mathbb{E}\left[u'(X+\eta)\right]}{u'\left(u^{-1}\left(\mathbb{E}\left[u(X+\eta)\right]\right)\right)} \\ &\leq -\beta + \alpha\frac{\mathbb{E}\left[u'(X+\eta)\right]}{u'\left(\mathbb{E}\left[X\right]+\eta\right)} \\ &\leq -\beta + \alpha\frac{u'(X_{\min}+\eta)}{u'\left(\mathbb{E}\left[X\right]+\eta\right)}. \end{aligned}$$

Hence the assumption on u implies that

$$\limsup_{\eta\uparrow\infty} H'(\eta) \leq \alpha - \beta < 0,$$

i.e., H is strictly decreasing for sufficiently large $\eta > 0$. In view of (i) this proves (ii).

(iii) If $u(0) = -\infty$ and $P(X = X_{\min}) > 0$ then $H(-X_{\min}) = -\infty$, so $\eta^* > -X_{\min}$. This completes the proof. □

Before we address concrete specifications of OEU risk measures in Example 4.10 below, we briefly investigate the dependence of η^* on the subjective time preference parameter α. It is straightforward to check that, ceteris paribus, η^* is an increasing function of the investor's time preference α; see also Vinel and Krokhmal [133, Proposition 6]:

4.8 Lemma. *If* $\alpha_1 < \alpha_2$*, then* $\eta^*(\alpha_1) \leq \eta^*(\alpha_2)$

Proof. Clearly,

$$-\beta\eta^*(\alpha_2) + \alpha_1 u^{-1}\left(\mathbb{E}\left[u\left(X + \eta^*(\alpha_2)\right)\right]\right) \leq -\beta\eta^*(\alpha_1) + \alpha_1 u^{-1}\left(\mathbb{E}\left[u\left(X + \eta^*(\alpha_1)\right)\right]\right). \quad (4.6)$$

We deny the conclusion and suppose that $\eta^*(\alpha_1) > \eta^*(\alpha_2)$. Then, due to (4.6) and to the assumption that $\alpha_1 < \alpha_2$, we get

$$0 < \beta(\eta^*(\alpha_1) - \eta^*(\alpha_2)) < \alpha_2\left(u^{-1}\left(\mathbb{E}\left[u\left(X + \eta^*(\alpha_1)\right)\right]\right) - u^{-1}\left(\mathbb{E}\left[u\left(X + \eta^*(\alpha_2)\right)\right]\right)\right),$$

which implies

$$-\beta\eta^*(\alpha_2) + \alpha_2 u^{-1}\left(\mathbb{E}\left[u\left(X + \eta^*(\alpha_2)\right)\right]\right) < -\beta\eta^*(\alpha_1) + \alpha_2 u^{-1}\left(\mathbb{E}\left[u\left(X + \eta^*(\alpha_1)\right)\right]\right).$$

The above equation, however, contradicts the definition of $\eta^*(\alpha_2)$ which completes the proof. □

We note that the bigger α, the more capital $\eta^*(\alpha)$ the investor decides to borrow at present to hedge against the risk of X. This behavior is intuitively understandable since bigger values of α give the investor a greater reward from the investment in $X + \eta^*(\alpha)$. In other words, if the investor attributes less value to the current possession of cash, then he is consequently more aimed at borrowing risk capital at present and investing in the future payoff $X + \eta^*(\alpha)$.

The extreme cases are covered by the following result:

4.9 Proposition. *Suppose that* $u \in \mathcal{U}_{DARA}$*. Then,*

(i) $\eta^* = -X_{\min}$ *if* $\alpha = 0$*, and*

(ii) $\rho^u(X) = -H(\infty) = -\lim_{\eta\uparrow\infty} H(\eta)$ *if* $\alpha = \beta$*.*

Proof. For $\eta > X_{\min}$ we have

$$H'(\eta) = -\beta + \alpha\frac{\mathbb{E}\left[u'(X + \eta)\right]}{u'\left(u^{-1}\left(\mathbb{E}\left[u(X + \eta)\right]\right)\right)}.$$

If $\alpha = 0$ it follows that $H'(\eta) = -\beta$, and hence $\eta^* = -X_{\min}$. If $\alpha = \beta$ we obtain

$$H'(\eta) = -\beta + \beta\frac{\mathbb{E}\left[u'(X + \eta)\right]}{g\left(\mathbb{E}\left[u(X + \eta)\right]\right)},$$

where

$$g(x) := u'(u^{-1}(x)).$$

Note that, if $u \in \mathcal{U}_{\mathrm{DARA}}$, then

$$g'(x) = \frac{u''(u^{-1}(x))}{u'(u^{-1}(x))} = -\varrho_u(u^{-1}(x)) \quad \text{is increasing,}$$

so g is convex. By Jensen's inequality, we conclude

$$H'(\eta) \geq -\beta + \beta = 0$$

for any $\eta > -X_{\min}$ and consequently $\rho^u(X) = -\lim_{\eta\uparrow\infty} H(\eta)$. □

We now present examples of convex optimal expected utility risk measures.

4.10 Example. (a) For the logarithmic utility function $u(x) = \ln(x)$ with constant relative risk aversion $x\varrho_u(x) = 1$, Proposition 4.7 (ii) implies that η^* is attained in $[-X_{\min}, \infty)$ and

$$\rho^u(X) = - \max_{\eta > -X_{\min}} \left\{ -\beta\eta + \alpha \exp\left(\mathbb{E}\left[\ln\left(X+\eta\right)\right]\right) \right\}.$$

(b) For a CRRA utility function $u(x) = \frac{1}{1-\gamma}\left(x^{1-\gamma} - 1\right)$ with $\gamma > 0, \gamma \neq 1$, Proposition 4.7 (ii) applies to show that η^* is attained in $[-X_{\min}, \infty)$ and

$$\rho^u(X) = - \max_{\eta > -X_{\min}} \left\{ -\beta\eta + \alpha \mathbb{E}\left[(X+\eta)^{1-\gamma}\right]^{\frac{1}{1-\gamma}} \right\}.$$

(c) For $u(x) = x$, a direct computation of (OEU) yields:

$$\begin{aligned} \rho^u(X) &= - \sup_{\eta > -X_{\min}} \left\{ -\beta\eta + \alpha\mathbb{E}[X+\eta] \right\} \\ &= - \sup_{\eta > -X_{\min}} \left\{ (\alpha - \beta)\eta + \alpha\mathbb{E}[X] \right\} \\ &= (\alpha - \beta) X_{\min} - \alpha\mathbb{E}[X]. \end{aligned}$$

Thus (OEU) reduces to a combination of the payoff's worst-case value and its expected value; for $\alpha = \beta$ the discounted expected loss obtains.[2]

(d) For $u(x) = 0$, OEU equals the negative discounted worst-case value of X:

$$\rho^u(X) = - \sup_{\eta > -X_{\min}} \left\{ -\beta\eta \right\} = -\beta X_{\min}.$$

This characterizes extremely risk averse investors that refuse to accept any risk and consequently set aside the maximum potential loss.[3]

(e) For a CARA utility $u(x) = \frac{1}{\gamma}\left(1 - \exp(-\gamma x)\right)$ with $\gamma > 0$, we have

$$\begin{aligned} \rho^u(X) &= - \sup_{\eta\in\mathbb{R}} \left\{ -\beta\eta + \alpha(-\tfrac{1}{\gamma}) \ln\left(1 - \gamma\mathbb{E}\left[\tfrac{1}{\gamma}\left(1 - \exp(-\gamma(X+\eta))\right)\right]\right) \right\} \\ &= - \sup_{\eta\in\mathbb{R}} \left\{ -\beta\eta + \alpha(-\tfrac{1}{\gamma}) \ln\left(\mathbb{E}\left[\exp(-\gamma(X+\eta))\right]\right) \right\} \\ &= - \sup_{\eta\in\mathbb{R}} \left\{ -\beta\eta + \alpha(-\tfrac{1}{\gamma}) \ln\left(\exp(-\gamma\eta)\mathbb{E}\left[\exp(-\gamma X)\right]\right) \right\} \\ &= - \sup_{\eta\in\mathbb{R}} \left\{ -\beta\eta + \alpha\eta - \tfrac{\alpha}{\gamma} \ln\left(\mathbb{E}\left[\exp(-\gamma X)\right]\right) \right\} \\ &\overset{\eta^* = -X_{\min}}{=} (\alpha - \beta) X_{\min} + \alpha\tfrac{1}{\gamma} \ln(\mathbb{E}[\exp(-\gamma X)]). \end{aligned}$$

In particular, in the limit $\alpha \uparrow \beta$ we obtain the classical entropic risk measure. At the end of this section, we return to this in a more general context.

[2] The functions in (c) and (d) are not strictly concave and therefore not in $\mathcal{U}$. We consider these merely formally here and interpret them as the limits of the utility functions $u(x) = \frac{1}{1-\gamma}\left(x^{1-\gamma} - 1\right)$ for $\gamma \downarrow 0$ and $\gamma \uparrow \infty$, respectively.

[3] See Footnote 2.

(f) Let us consider piecewise linear functions of the form $u(x) = -\gamma_1 x^- + \gamma_2 x^+$, $\gamma_1 \geq 1 > \gamma_2 > 0$, $x^- := \max\{-x,0\}$, $x^+ := \max\{x,0\}$. Obviously, these functions do not fulfill the conditions of Definition 3.5. Nevertheless, they provide a convenient way of representing utilities. Proceeding formally, we have for all $X \in \mathcal{X}$:

$$\rho^u(X) = -\sup_{\eta\in\mathbb{R}}\Bigg\{-\beta\eta + \alpha\Bigg(-\frac{1}{\gamma_1}\left(\mathbb{E}\left[-\gamma_1(X+\eta)^- + \gamma_2(X+\eta)^+\right]\right)^- \\ + \frac{1}{\gamma_2}\left(\mathbb{E}\left[-\gamma_1(X+\eta)^- + \gamma_2(X+\eta)^+\right]\right)^+\Bigg)\Bigg\}$$

with the optimal solution

$$\eta^* = \begin{cases} -\beta F^{-1}\left(\frac{\frac{1}{\alpha}-\frac{\gamma_2}{\gamma_1}\frac{1}{\beta}}{1-\frac{\gamma_2}{\gamma_1}}\right), & \text{if } \gamma_1\mathbb{E}\left[(X+\eta)^-\right] - \gamma_2\mathbb{E}\left[(X+\eta)^+\right] > 0, \\ -\beta F^{-1}\left(\frac{\frac{1}{\alpha}-\frac{1}{\beta}}{\frac{\gamma_1}{\gamma_2}-1}\right), & \text{otherwise} \end{cases}, \tag{4.7}$$

where F is the distribution function of the random variable X. If we drop the assumption that γ_2 needs to be positive and set $\gamma_1 = 1$, $\gamma_2 = 0$, i.e., $u(x) = -(-x)^+$, $\alpha = \frac{1}{\lambda}$, $\beta = 1$, we get

$$\rho^u(X) = \inf_{\eta\in\mathbb{R}}\left\{\eta + \frac{1}{\lambda}\mathbb{E}\left[(-X-\eta)^+\right]\right\} = \begin{cases} -\infty, \text{ for } \lambda > 1, \\ \mathbb{E}[-X], \text{ for } \lambda = 1, \\ AV@R^\lambda(X), \text{ for } \lambda < 1, \end{cases}$$

where the identity for $\lambda < 1$ is due to [117, Theorem 1]. In this case, (4.7) simplifies to

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

which is another way to derive $AV@R$ for $\lambda < 1$:

$$\begin{aligned} \rho^u(X) &= -F^{-1}(\lambda) + \frac{1}{\lambda}\int_{-\infty}^{F^{-1}(\lambda)} x dF(x) \\ &= \frac{1}{\lambda}\int_0^\lambda F^{-1}(\kappa)d\kappa \\ &= \frac{1}{\lambda}\int_0^\lambda V@R^\kappa(X)d\kappa \\ &= AV@R^\lambda(X). \end{aligned}$$

For $0 < \lambda < 1$, $F^{-1}(\lambda)$ is the λ-quantile of X.

It follows that we can formally regard $AV@R$ as a special case of OEU with a (piecewise) linear utility function and $\alpha > \beta$. From an OEU perspective, this is a degenerate case; in the more natural case $\alpha < \beta$, it follows that $\rho^u(X) = -\infty$ which is reasonable since u induces the investor to spend ∞ dollars today knowing that less money needs to be payed back.

Note that the utility functions of (a) and (b) can be written as one single function:

$$u(x) = \begin{cases} \frac{1}{1-\gamma}\left(x^{1-\gamma} - 1\right), \; \gamma > 0, \gamma \neq 1 \\ \ln(x), \; \gamma = 1 \end{cases},$$

since, by virtue of l'Hôpital's rule, logarithmic utility is the limiting case of power utility for $\gamma \to 1$:

$$\lim_{\gamma \to 1} \frac{x^{1-\gamma} - 1}{1 - \gamma} = \lim_{\gamma \to 1} \frac{-x^{1-\gamma} \ln(x)}{-1} = \ln(x).$$

4.3 Further properties and comparative statistics

Given two utility functions $u, v \in \mathcal{U}$ and a financial position $X \in \mathcal{X}$, it is natural to ask how $\rho^u(X)$ relates to $\rho^v(X)$. The following result provides an answer to this question in terms of the "degree of concavity" of u and v.

4.11 Proposition. *Let $u, v \in \mathcal{U}$. If u is more concave than v, i.e., if there exists a strictly increasing concave function g such that $u = g \circ v$, then*

$$\rho^u(X) \geq \rho^v(X).$$

Proof. Proposition 3.8 implies that u is more concave than v if and only if

$$u^{-1}\left(\mathbb{E}\left[u(X)\right]\right) \leq v^{-1}\left(\mathbb{E}\left[v(X)\right]\right)$$

for all $X \in \mathcal{X}$. It follows that

$$-\beta\eta + \alpha u^{-1}\left(\mathbb{E}\left[u(X + \eta)\right]\right) \leq -\beta\eta + \alpha v^{-1}\left(\mathbb{E}\left[v(X + \eta)\right]\right) \quad \text{for all } \eta \in \mathbb{R},$$

and taking the supremum on both sides yields $\rho^u(X) \geq \rho^v(X)$. □

Proposition 4.11 is intuitively obvious: An investor with a more concave utility function is more risk averse, hence requires more reserve capital than an investor with a less concave utility function. Combining Example 4.10 (c), (d) with the argument in the proof of Proposition 4.11, we obtain the following general upper and lower bounds for OEU:

4.12 Corollary. *Suppose that $u \in \mathcal{U}$. Then for all $X \in \mathcal{X}$ we have*

$$-\beta X_{\min} - \alpha(\mathbb{E}[X] - X_{\min}) \leq \rho^u(X) \leq -\beta X_{\min}.$$

For positive payoffs, we obtain an alternative upper bound for OEU in terms of the certainty equivalent:

4.13 Proposition. *If $X \in \mathcal{X}$ is such that $X_{\min} \geq 0$, then*

$$\rho^u(X) \leq -\alpha u^{-1}\left(\mathbb{E}\left[u\left(X\right)\right]\right).$$

Proof. If $X_{\min} \geq 0$ then $\eta \in (-X_{\min}, \infty)$ for all $\eta > 0$, and thus

$$\rho^u(X) \leq -\lim_{\eta \downarrow 0} H(\eta) = -H(0) = -\alpha u^{-1}\left(\mathbb{E}\left[u\left(X\right)\right]\right)$$

using the continuity established in Lemma 4.5. □

We wish to emphasize that Proposition 4.13 applies only to positive payoffs as in the typical asset management applications; the other important case of positions that may lead to negative payoffs is not covered by this result, so it is not true that ρ^u is less conservative than ρ^{CE}.

Next, we return to the case $\alpha = \beta$ and identify the limit in Proposition 4.9 (ii) in terms of the utility function's asymptotic absolute risk aversion. More precisely, if $\varrho_u(x) \to \gamma$ as $x \uparrow \infty$ and $\alpha = \beta$, then OEU reduces to the trivial risk measures in Examples 4.10 (c) and (e).

4.14 Proposition. *Let $\alpha = \beta$ and $u \in \mathcal{U}_{DARA}$ and set $\gamma := \lim_{x \uparrow \infty} \varrho_u(x)$. Then*

$$\rho^u(X) = \begin{cases} \beta \frac{1}{\gamma} \ln\left(\mathbb{E}\left[\exp(-\gamma X)\right]\right), & \gamma > 0 \\ \beta \mathbb{E}\left[-X\right], & \gamma = 0 \end{cases}.$$

Proof. Define

$$u_\gamma(x) := \begin{cases} \frac{1}{\gamma}\left(1 - \exp(-\gamma x)\right), & \gamma > 0 \\ x, & \gamma = 0 \end{cases}.$$

Proposition 4.9 implies that

$$\rho^u(X) = -\lim_{\eta \uparrow \infty} H(\eta).$$

Since u has decreasing absolute risk aversion and $\lim_{x \uparrow \infty} \varrho_u(x) = \gamma$, whereas u_γ has constant absolute risk aversion γ, it follows that for all $\eta \in \mathbb{R}$ we have

$$\varrho_u(X + \eta) \geq \varrho_{u_\gamma}(X + \eta) = \gamma.$$

On the other hand, for all $\gamma_0 > \gamma$ there exists some $\eta_0 \in \mathbb{R}$ such that for all $\eta > \eta_0$

$$\varrho_u(X + \eta) \leq \varrho_{u_{\gamma_0}}(X + \eta) = \gamma_0.$$

Thus, following [114, Theorem 1], we have for all $\gamma_0 > \gamma$

$$\lim_{\eta \uparrow \infty}\left\{u_{\gamma_0}^{-1}\left(\mathbb{E}\left[u_{\gamma_0}(X + \eta)\right]\right)\right\} \leq \lim_{\eta \uparrow \infty}\left\{u^{-1}\left(\mathbb{E}\left[u(X + \eta)\right]\right)\right\} \leq \lim_{\eta \uparrow \infty}\left\{u_{\gamma}^{-1}\left(\mathbb{E}\left[u_\gamma(X + \eta)\right]\right)\right\}.$$

It follows that

$$\lim_{\eta \uparrow \infty} H(\eta) = \lim_{\eta \uparrow \infty}\left\{-\beta\eta + \alpha u_\gamma^{-1}\left(\mathbb{E}\left[u_\gamma(X + \eta)\right]\right)\right\} \overset{\text{Ex. 4.10 (c), (e)}}{=} \begin{cases} -\beta \frac{1}{\gamma} \ln\left(\mathbb{E}\left[\exp(-\gamma X)\right]\right), & \gamma > 0 \\ -\beta \mathbb{E}\left[-X\right], & \gamma = 0 \end{cases}.$$

□

The risk measure that appears in Proposition 4.14 is a variant of the well-known *entropic risk measure*. It is obtained as a special case of OEU if we set $\alpha = \beta$ and $u(x) = \frac{1}{\gamma}(1 - \exp(-\gamma x))$ for some $\gamma > 0$:

$$\rho^u(\beta X) = \beta \frac{1}{\gamma} \ln(\mathbb{E}[\exp(-\gamma \beta X)]) = \beta \rho^{\text{ent}}(X).$$

The entropic risk measure is introduced in [67] and [74]; for further information on ρ^{ent} we refer to Föllmer and Knispel [65] and the references therein.

Finally, we address the ranking of payoffs by OEU. If $X \in \mathcal{X}$ first-order stochastically dominates $Y \in \mathcal{X}$, then

$$\mathbb{E}[v(X)] \geq \mathbb{E}[v(Y)] \quad \text{for every increasing function } v$$

and thus $\rho^u(X) \leq \rho^u(Y)$. More generally, X second-order stochastically dominates Y if

$$\mathbb{E}[v(X)] \geq \mathbb{E}[v(Y)] \quad \text{for every concave increasing function } v.$$

The following result shows that OEU also preserves second-order stochastic dominance.

4.15 Proposition. *Suppose $X,Y \in \mathcal{X}$ are such that*

$$\mathbb{E}[v(X)] \geq \mathbb{E}[v(Y)]$$

for all increasing, concave functions $v : \mathbb{R} \to [-\infty,\infty)$. Then we have $\rho^u(X) \leq \rho^u(Y)$ for all $u \in \mathcal{U}$.

Proof. Let $u \in \mathcal{U}$ and set

$$H(\eta,X) := -\beta\eta + \alpha u^{-1}(\mathbb{E}[u(X+\eta)]).$$

For any $\eta \in \mathbb{R}$, define $v_\eta(x) := u(x+\eta)$, $x \in \mathbb{R}_+$. Then, by assumption, we have

$$\alpha u^{-1}(\mathbb{E}[v_\eta(X)]) \geq \alpha u^{-1}(\mathbb{E}[v_\eta(Y)])$$

and it follows that

$$H(\eta,X) \geq H(\eta,Y).$$

This applies for every $\eta \in \mathbb{R}$; consequently

$$\rho^u(X) = -\sup_{\eta \in \mathbb{R}} H(\eta,X) \leq -\sup_{\eta \in \mathbb{R}} H(\eta,X) = \rho^u(Y)$$

and the proof is complete. □

4.4 Homogeneity of OEU

The following theorem shows how homogeneity of ρ^u depends on the relative risk aversion of $u \in \mathcal{U}$. The theorem particularly specifies when ρ^u is a coherent risk measure. Recall that it is impossible for $u \in \mathcal{U}$ to have decreasing relative risk aversion.

4.16 Theorem. *Suppose that $u \in \mathcal{U}_{IRRA}$ and $X \in \mathcal{X}$. Then,*

(i) $\rho^u(\lambda X) \leq \lambda\rho^u(X)$ *for all* $\lambda \in [0,1]$, *and*

(ii) $\rho^u(\lambda X) \geq \lambda\rho^u(X)$ *for* $\lambda \in [1,\infty)$.

If $u \in \mathcal{U}_{CRRA}$, then

$$\rho^u(\lambda X) = \lambda\rho^u(X) \quad \textit{for all } \lambda \geq 0.$$

Proof. ρ^u is a convex risk measure by Theorem 4.3. In particular, for $\lambda = 0$ we have $\rho^u(\lambda X) = \lambda\rho^u(X)$ by (N). Since by [114, Theorem 6] for $u \in \mathcal{U}_{\mathrm{IRRA}}$ it follows that $\frac{1}{\lambda}u^{-1}(\mathbb{E}[u(\lambda X)])$ is a decreasing function of $\lambda > 0$, we conclude that

$$\tfrac{1}{\lambda}\left(-\beta\eta + \alpha u^{-1}\left(\mathbb{E}\left[u(\lambda X + \eta)\right]\right)\right)$$

is decreasing in λ for all $\eta \in \mathbb{R}$. This implies that

$$-\tfrac{1}{\lambda}\sup_{\eta\in\mathbb{R}}\left\{-\beta\eta + \alpha u^{-1}\left(\mathbb{E}\left[u\left(\lambda X + \eta\right)\right]\right)\right\}$$

is increasing in λ, and therefore

$$\begin{aligned}
\rho^u(\lambda X) &\leq \lambda\rho^u(X), \quad 0 \leq \lambda < 1,\\
\rho^u(\lambda X) &= \lambda\rho^u(X), \quad \lambda = 1,\\
\rho^u(\lambda X) &\geq \lambda\rho^u(X), \quad \lambda > 1.
\end{aligned}$$

By [114, Theorem 6] the same equivalence holds if "increasing" is replaced by "decreasing". Hence ρ^u is positively homogeneous for $u \in \mathcal{U}_{\mathrm{CRRA}}$. □

Recall that, as pointed out in Section 3.2, constant (or slightly increasing) relative risk aversion is a reasonable assumption for real-world investors.

4.17 Corollary. *ρ^u is a coherent risk measure for logarithmic and power utility, and for the linear utility functions in Examples 4.10 (c) and (d).*

To the best of our knowledge OEU is the only existing utility-based risk measure that is coherent for power utility functions (ρ^{CE} is not even convex, ρ^{OCE} is not coherent, and ρ^{SR} reduces to a worst-case risk measure for CRRA utility functions).

4.5 Recoverability

This section deals with the problem of recovering the utility function u from a given risk measure ρ^u. For fixed $x_1,x_2 \in \mathbb{R}$ with $x_2 < x_1$ we introduce the random variable

$$\hat{X}_p := \begin{cases} x_1, & \text{with probability } 1-p, \\ x_2, & \text{with probability } p, \end{cases}$$

and the related function

$$H(\eta,p) := -\beta\eta + \alpha u^{-1}\left((1-p)u(x_1+\eta) + pu(x_2+\eta)\right) \quad \text{for } p \in [0,1].$$

The following lemma shows that $H(\eta,p)$ converges uniformly to $H(\eta,1)$ if p goes to 1. This result will become particularly useful since it also provides a limit of OEU if p tends to the boundary value 1 for random variables $\hat{X}_p$.

4.18 Lemma. *For any $u \in \mathcal{U}$*

$$H(\eta,p) \xrightarrow[unif.]{p\uparrow 1} H(\eta,1) = -\beta\eta + \alpha(x_2+\eta).$$

Proof. Since u is assumed to be concave, by Jensen's inequality

$$u^{-1}\left(\mathbb{E}\left[u\left(\hat{X}_p+\eta\right)\right]\right) \le \mathbb{E}\left[\hat{X}_p+\eta\right]$$

for all $\eta \in \mathbb{R}$ and consequently

$$\underbrace{u^{-1}\left(\mathbb{E}\left[u\left(\hat{X}_p+\eta\right)\right]\right)}_{\ge \eta + x_2} - (\eta + x_2) \le \underbrace{(1-p)(x_1-x_2)}_{\ge 0}$$

for all $\eta \ge -x_2$. Thus,

$$\begin{aligned} \sup_{\eta > -x_2} |H(\eta,p) - H(\eta,1)| &= \alpha \sup_{\eta > -x_2} \left|u^{-1}\left(\mathbb{E}\left[u\left(\hat{X}_p+\eta\right)\right]\right) - (\eta + x_2)\right| \\ &\le \alpha \sup_{\eta > -x_2} |(1-p)(x_1-x_2)| \\ &\xrightarrow{p\uparrow 1} 0. \end{aligned}$$

□

The following result shows how to regain the utility function from a given optimal expected utility risk measure.

4.19 Proposition. *If $u \in \mathcal{U}$ is such that $u(0) = 0$, $u'(0) := \lim_{x\downarrow 0} u'(x) = 1$, then*

$$u(x) = \frac{1}{\alpha} \lim_{p\uparrow 1} \left(\frac{\partial}{\partial p} \rho^u(\hat{X}_p^0)\right), \quad \text{with } \hat{X}_p^0 := \begin{cases} x, & \text{with probability } 1-p, \quad x > 0, \\ 0, & \text{with probability } p \end{cases}.$$

Proof. Firstly, let us consider

$$\rho^u(\hat{X}_p^0) = -\sup_{\eta\geq 0} H(\eta,p) = -H(\eta^*,p),$$

where, due to Proposition 4.7, $\eta^* = \eta^*(p)$ is the value which optimizes (OEU). With Lemma 4.18 it holds that

$$H(\eta,p) \xrightarrow[\text{unif.}]{p\uparrow 1} H(\eta,1) = (\alpha-\beta)\eta \ < 0 \quad \text{for all } \eta > 0,$$

which implies that there exists some $p_f < 1$ such that

$$H(\eta,p) < 0 \quad \text{for all } \eta > 0 \text{ for all } p_f \leq p \leq 1.$$

This means that $\eta^* = 0$ for all $p_f \leq p \leq 1$ since $H(0,p) = \alpha u^{-1}\left((1-p)u(x)\right) > 0$ if $u(0) = 0$. The partial derivative of $\rho^u(\hat{X}_p^0)$ with respect to p for $p_f \leq p \leq 1$ is

$$\begin{aligned}\frac{\partial}{\partial p}\rho^u(\hat{X}_p^0) &= \frac{\partial}{\partial p}\left(-\alpha u^{-1}\left((1-p)u(x)+pu(0)\right)\right)\\ &= -\alpha\frac{-u(x)+u(0)}{u'\left(u^{-1}((1-p)u(x)+pu(0))\right)}\end{aligned}$$

and, due to $u(0) = 0$, $u^{-1}(0) = 0$, $u'(0) = 1$, we have

$$\frac{1}{\alpha}\lim_{p\uparrow 1}\left(\frac{\partial}{\partial p}\rho^u(\hat{X}_p^0)\right) = \frac{u(x)-u(0)}{u'(0)} = u(x),$$

which concludes the proof. □

If $u \in \mathcal{U}$ such that $u(0) \neq 0$ or $u'(0) \neq 1$, and u satisfies the condition from Proposition 4.7 (ii), then, for any $x > 0$ and the related $\hat{X}_p^0$, (OEU) is optimized by some $\eta_x^* \in [0,\infty)$. Thus, we can apply Proposition 4.19 for the normalized utility function $\tilde{u}$ which is related to u through

$$\tilde{u}(x) = \frac{u(x+\eta_x^*)-u(\eta_x^*)}{u'(\eta_x^*)}.$$

Obviously, $\tilde{u} \in \mathcal{U}$ if $u \in \mathcal{U}$ and therefore, given any $u \in \mathcal{U}$, Proposition 4.19 holds true for the normalized counterpart $\tilde{u}$. We see that the normalized utility function $\tilde{u}$ evaluated at x is the compounded ($\frac{1}{\alpha}$) infinitesimal change of the risk of $\hat{X}_p^0$ attaining x. In particular, the utility function is, up to affine transformations, uniquely determined by ρ^u.

We now make use of Proposition 4.19 to recover u from OEU of Example 4.10 (e).

4.20 Example. Let $\rho^u(X) = (\alpha-\beta)X_{\min} + \alpha\frac{1}{\gamma}\ln\left(\mathbb{E}\left[\exp(-\gamma X)\right]\right)$. Then,

$$\frac{\partial}{\partial p}\rho^u(\hat{X}_p^0) = \frac{\partial}{\partial p}\left(\alpha\frac{1}{\gamma}\ln\left((1-p)\exp(-\gamma x)+p\right)\right) = \alpha\frac{1}{\gamma}\frac{-\exp(-\gamma x)+1}{(1-p)\exp(-\gamma x)+p}$$

which leads to the utility function

$$u(x) = \frac{1}{\alpha}\lim_{p\uparrow 1}\left(\frac{\partial}{\partial p}\rho^u(\hat{X}_p^0)\right) = \frac{1}{\gamma}\left(1-\exp(-\gamma x)\right).$$

4.6 Dual representation of OEU

This section includes the dual representation of OEU, where we mainly follow the textbook of Föllmer and Schied [69] for the theory of *robust representation of convex risk measures*. Our aim is to point out how the penalty function of the dual representation of OEU can be understood as a *supremal convolution* of a (generalized) relative entropy. Note that in this section we consider sequences of random variables $\{X^{(n)}\}$ and write

$$X^{(n)}_{\min} := \text{ess inf } X^{(n)}$$

for the left support of any $X^{(n)} \in \mathcal{X}$.

4.21 Remark. Throughout this section we make the technical assumptions that $(\Omega,\mathcal{F},P)$ does not have atoms and that $L^2(\Omega,\mathcal{F},P)$ is separable. Under these assumptions, Jouini, Schachermayer and Touzi [88, Theorem 2.1] have shown that law-invariance of a convex risk measure on $L^\infty(\Omega,\mathcal{F},P)$ implies continuity from above. Thus optimal expected utility risk measures are continuous from above, i.e., for any $u \in \mathcal{U}$ and any sequence $\{X^{(n)}\} \subset \mathcal{X}$ with $X^{(n)} \downarrow X$ for some $X \in \mathcal{X}$, we have

$$\rho^u(X^{(n)}) \uparrow \rho^u(X).$$

If, however, $\{X^{(n)}\}$ is a sequence that increases pointwise to $X \in \mathcal{X}$, but

$$\lim_{n\to\infty} X^{(n)}_{\min} < X_{\min},$$

i.e., there is a positive probability that $\lim_{n\to\infty} X^{(n)}$ has a another worst-case scenario than X, then OEU typically captures this additional risk. This is the reason why ρ^u is in general not continuous from below which is illustrated by the following example: Consider any $X \in \mathcal{X}$ and the sequence $\{X^{(n)}\} \subset \mathcal{X}$ given by

$$X^{(n)} := X - \mathbf{1}_{[0,\frac{1}{n}]},$$

where we choose the probability space $([0,1],\mathcal{B}([0,1]),\lambda)$ consisting of the real interval $[0,1]$ as the sample space, the Borel σ-algebra on $[0,1]$ (characterized as the minimal σ-field generated by the open intervals (a,b) on $[0,1]$) and, for any $B \in \mathcal{B}([0,1])$, the Lebesgue measure $\lambda(\cdot)$ defined as the sum of the lengths of the intervals contained in B. Note that $X^{(n)} \uparrow X$. Due to monotonicity of ρ^u,

$$\begin{aligned}
\inf_{n\in\mathbb{R}} \rho^u\left(X^{(n)}\right) &= \inf_{\eta\in\mathbb{R}}\inf_{n\in\mathbb{R}} \left\{\beta\eta - \alpha u^{-1}\left(\mathbb{E}\left[u\left(X^{(n)}+\eta\right)\right]\right)\right\} \\
&\geq \inf_{\eta\in\mathbb{R}} \left\{\beta\eta - \alpha u^{-1}\left(\mathbb{E}\left[u\left(X+\eta\right)\right]\right)\right\} \\
&= \rho^u(X),
\end{aligned}$$

where the inequality is due to the fact that for $\eta \in [-X_{\min}, -X_{\min}+1)$, $X^{(n)}+\eta$ attains strictly negative values for all $n \in \mathbb{R}$, while $X+\eta$ is positive. Now, if we choose $u \in \mathcal{U}$ such that the infimum in

$$\rho^u(X) = \inf_{\eta\in\mathbb{R}} \left\{\beta\eta - \alpha u^{-1}\left(\mathbb{E}\left[u\left(X+\eta\right)\right]\right)\right\}$$

is attained uniquely at $\eta^* \in [-X_{\min}, -X_{\min}+1)$, the above inequality is strict and we have an example for OEU not being continuous from below.

Since ρ^u is continuous from above, for $\mathcal{X} = L^\infty(\Omega,\mathcal{F},P)$ it admits the dual representation

$$\rho(X) = \sup_{Q\in\mathcal{M}_1(P)} (\mathbb{E}_Q[-\beta X] - \vartheta_{\min}(Q)),$$

due to Theorem 2.7, where $\vartheta_{\min}$ is the minimal penalty function. A formula for $\vartheta_{\min}$ is given by [69, Remark 4.17 (a)]

$$\vartheta_{\min}(Q) = \sup_{X\in\mathcal{X}} \left\{\mathbb{E}^Q[-\beta X] - \rho(X)\right\}, \tag{4.8}$$

i.e., for any probabilistic model $Q \in \mathcal{M}_1(P)$ we compute $\vartheta_{\min}$ as the worst case of the expected value of $-\beta X$ under Q, but reduced by $\rho(X)$. Note that if ρ is a utility-based risk measure, it is computed under the given probability measure P, "the one which is taken most seriously" ([69, page 201]). From Remark 2.8 we know that the relative entropy of any probability measure Q with respect to P is

$$H(Q \mid P) = \sup_{X\in\mathcal{X}} \left\{\mathbb{E}^Q[-\beta X] - \ln\left(\mathbb{E}\left[\exp(-\beta X)\right]\right)\right\}$$

which justifies the conclusion that the minimal penalty function (4.8) for the robust representation of OEU for

$$u(x) = \frac{1}{\gamma}\left(1 - \exp(-\gamma x)\right),\ \gamma > 0,$$

is just the relative entropy (up to an additive constant). Consequently, the (discounted) negative certainty equivalent

$$\rho_\alpha^{\mathrm{CE}}(X) = -\alpha u^{-1}\left(\mathbb{E}\left[u(X)\right]\right)$$

generalizes the relative entropy since it has the minimal penalty function

$$\vartheta_{\min}^{\mathrm{CE}}(Q) := \sup_{X\in\mathcal{X}} \left\{\mathbb{E}^Q[-\beta X] + \alpha u^{-1}\left(\mathbb{E}\left[u(X)\right]\right)\right\}.$$

The optimal expected utility risk measure ρ^u includes the supremal convolution to this generalization of the relative entropy:

$$\begin{aligned}
\vartheta_{\min}^{u}(Q) &:= \sup_{X\in\mathcal{X}} \left\{\mathbb{E}^Q[-\beta X] + \sup_{\eta\in\mathbb{R}} \left\{-\beta\eta + \alpha u^{-1}\left(\mathbb{E}\left[u\left(X+\eta\right)\right]\right)\right\}\right\} \\
&= \sup_{X\in\mathcal{X}} \sup_{\eta\in\mathbb{R}} \left\{\mathbb{E}^Q[-\beta X] - \beta\eta + \alpha u^{-1}\left(\mathbb{E}\left[u\left(X+\eta\right)\right]\right)\right\} \\
&= \sup_{X\in\mathcal{X}} \sup_{\eta\in\mathbb{R}} \left\{\mathbb{E}^Q[-\beta(X+\eta)] + \alpha u^{-1}\left(\mathbb{E}\left[u\left(X+\eta\right)\right]\right)\right\} \\
&= \sup_{\eta\in\mathbb{R}} \left\{\sup_{X_\eta\in\mathcal{X}} \left\{\mathbb{E}^Q[-\beta X_\eta] + \alpha u^{-1}\left(\mathbb{E}\left[u\left(X_\eta\right)\right]\right)\right\}\right\},
\end{aligned}$$

with $X_\eta := X + \eta$.

4.7 Applications of OEU

In this section we illustrate and compare OEU, $V@R$ and $AV@R$ using simplified real-world examples of risk assessment. Further, we consider a one-period optimal investment problem for OEU.

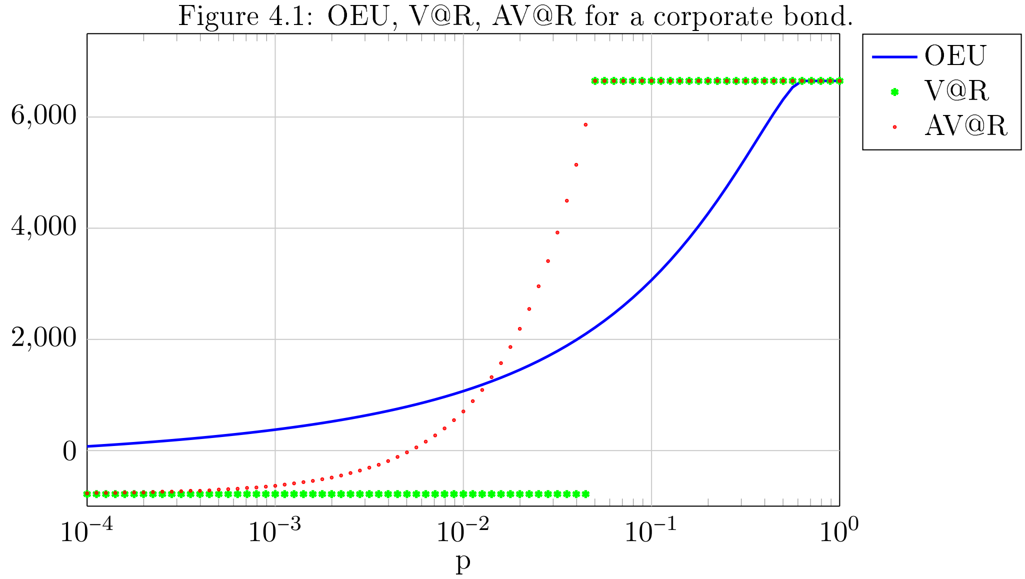
4.7.1 Comparison of OEU with other risk measures

4.22 Example. An investor wants to put 10,000 in a corporate bond with maturity 4 years and an annual interest rate of 2%. The estimated probability of default of the corporate within the next 4 years is p. In the event of solvency the estimated loss given default is 70%. Thus the investor faces an investment opportunity with two scenarios: He either makes a profit of $10{,}000 \cdot (1.02^4 - 1) = 824$ or he loses 7,000. The corresponding financial position is

$$X = \begin{cases} 824 \text{ with } 1-p, \\ -7{,}000 \text{ with } p. \end{cases}$$

The investor has to decide how to evaluate this investment opportunity. $V@R$ does not turn out to be a good choice in this setting: Depending on p, $V@R$ is either $-\beta \cdot 824$ or $\beta \cdot 7{,}000$, but never attains strictly intermediate values. $AV@R$ is $\beta \cdot 7{,}000$ for $p > \lambda$ and tends to $-\beta \cdot 824$ for $p \downarrow 0$. However, $AV@R$ is mostly close to $-\beta \cdot 824$ or $\beta \cdot 7{,}000$. Here, we choose $\beta = 0.95$, $\lambda = 0.05$. On the other hand, ρ^u is a risk measure that is easy to interpret and attains reasonable values for any choice of p: Consider an investor with a power utility function $u(x) = \frac{1}{1-\gamma}\left(x^{1-\gamma} - 1\right)$ with $\gamma = 10$ and subjective rate of time preference $\alpha = 0.9$. Figure 4.1 illustrates these three risk measures as functions of the corporate's probability of default p.

Figure 4.1: OEU, V@R, AV@R for a corporate bond.

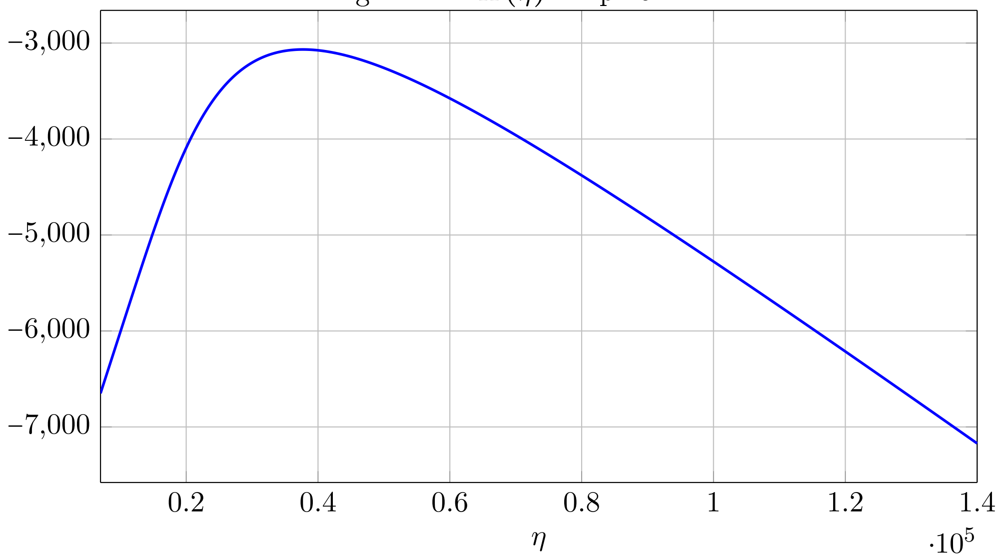


The corresponding function $H(\eta)$ for $p = 0.1$ is illustrated in Figure 4.2. Note that the supremum in (OEU) for $p = 0.1$ is attained at $\eta^* = 37{,}718$ and $\rho^u(X, 0.1) = 3{,}067$.

4.23 Remark. Note that for

$$\hat{X}_p = \begin{cases} x_1, & \text{with } 1-p, \\ x_2, & \text{with } p, \end{cases} \qquad x_1, x_2 \in \mathbb{R},\ x_2 < x_1,$$

Figure 4.2: $H(\eta)$ for p=0.1.



as we can see in Figure 4.1,

$$\rho^u(\hat{X}_p) \xrightarrow{p\downarrow 0} (\alpha - \beta)(x_2 - x_1) - \beta x_1,$$

whereas $V@R^\lambda(\hat{X}_p)$, $AV@R^\lambda(\hat{X}_p) \xrightarrow{p\downarrow 0} -\beta x_1$. The difference $-(\beta - \alpha)(x_2 - x_1)$ can be interpreted as follows: As long as X can attain x_2 with some positive probability $p > 0$, OEU takes into account this "worst-case scenario", and the investor sets aside a suitable amount of risk capital. Note that in the special case $\alpha = \beta$, OEU also yields

$$\rho^u(\hat{X}_p) \xrightarrow{p\downarrow 0} -\beta x_1.$$

On the other hand, if $p = 0$, i.e., the financial position X_p only consists of the certain payout x_1, then due to Remark 4.4 we get

$$\rho^u(\hat{X}_p) = -\beta x_1.$$

This highlights that, while OEU in general depends on the entire distribution of the payoff, it is sensitive to rare, yet bad scenarios.

4.24 Example. Following Giesecke, Schmidt and Weber [78, Example 3.3], we modify the payoff of the previous example by adding another possible payout x_3 to $\hat{X}_p$. Formally, we consider the financial position

$$X_p^3 = \begin{cases} x_1, & \text{with } 1 - p, \\ x_2, & \text{with } \frac{p}{2}, \\ x_3, & \text{with } \frac{p}{2}, \end{cases} \qquad x_1, x_2, x_3 \in \mathbb{R},\ x_3 < x_2 < x_1.$$

Then we have

$$V@R^p(X_p^3) = -\beta x_1$$

for any choice of x_2 and x_3 as well as

$$AV@R^p(X_p^3) = -\beta\frac{x_2 + x_3}{2}.$$

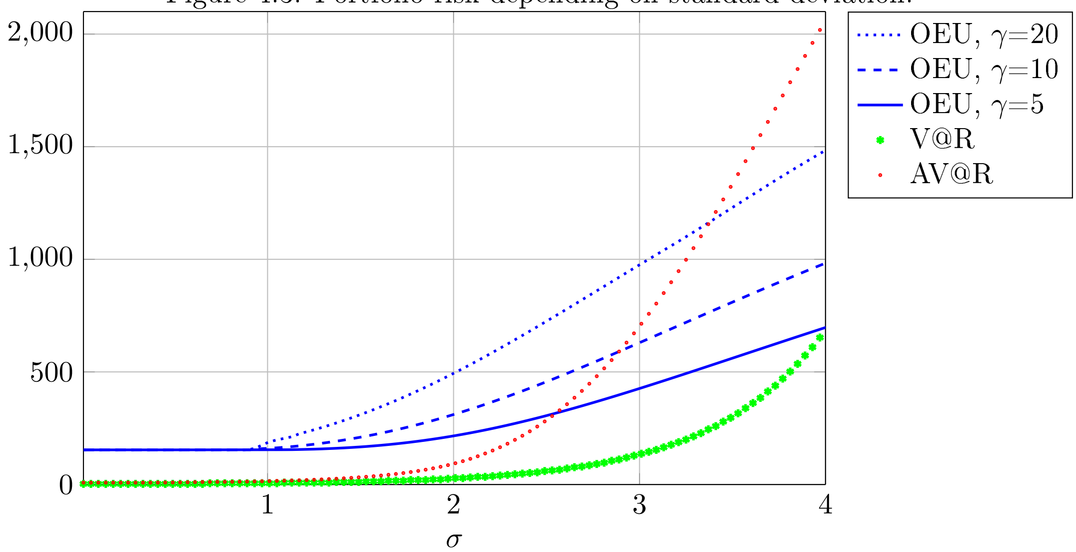
Thus $V@R$ is not suitable to measure the potential downside risk of X_p^3, and even $AV@R$ is insensitive to changes of X_p^3 as long as $x_2 + x_3$ remains constant. OEU, on the other hand, is non-constant on any hyperplane in the space (x_1,x_2,x_3).

4.25 Example. Let $-X$ be the portfolio loss of a financial institution which we assume to be floored log-normally distributed:

$$X = -(Z \wedge 3{,}000), \quad \text{where} \quad Z \sim \ln\mathcal{N}(0,\sigma),\ \sigma \in (0,4).$$

We wish to compare the sensitivities of $V@R$, $AV@R$ and ρ^u with respect to changes in the volatility parameter σ. We choose the level $\lambda = 0.05$, ρ^u is evaluated with $u(x) = \frac{1}{1-\gamma}\left(x^{1-\gamma} - 1\right)$ for $\alpha = 0.9$, $\beta = 0.95$, and $\gamma = 5,10,20$. Figure 4.3 illustrates the associated risk measures as functions of σ.

Figure 4.3: Portfolio risk depending on standard deviation.



OEU, $V@R$ and $AV@R$ all appear suitable for detecting risks due to heavy tails in the portfolio distribution.

4.7.2 One-period investment in one safe asset and in one stock

This section deals with an application of the functional

$$S^u(X) := -\rho^u(X) = \sup_{\eta\in\mathbb{R}}\left\{-\beta\eta + \alpha u^{-1}\left(\mathbb{E}\left[u\left(X+\eta\right)\right]\right)\right\}$$

to a model of investment in a risky/safe pair of assets in a single period.

To this end, we consider an investor with utility function $u \in \mathcal{U}$ and initial wealth $W_0 > 0$ and assume that there is a riskless security that pays a rate of return equal to $R_f \geq 1$ and one stock to invest in that has the initial price $x_0 > 0$ and at the end of the period has the uncertain price

$$\hat{X}_p = \begin{cases} x_1, & \text{with } 1-p, \\ x_2, & \text{with } p \end{cases}, \quad x_1 > x_0 > x_2 \geq 0$$

such that the return of the risky asset is

$$R = \begin{cases} R_1 = \frac{x_1}{x_0}, & \text{with } 1-p, \\ R_2 = \frac{x_2}{x_0}, & \text{with } p \end{cases}, \quad R_1 > R_f > R_2 \geq 0.$$

The dollar amount that the investor puts in the stock at the beginning of the period is w, thus $W_0 - w$ is the initial investment in the riskless security. We consider here the existence of short-sale constraints, i.e., w is restricted to lie in between 0 and W_0. Hence, the investor's wealth at the end of the considered period is

$$W_1 = R_f(W_0 - w) + Rw, \quad 0 \leq w \leq W_0.$$

A similar model can be found, amongst others, in Arrow [6, Chapter 3] and in [81, Chapter 5]. Such a model is usually analyzed via the maximal expected utility principle where the optimal investment w^* is the solution of

$$\sup_{0 \leq w \leq W_0} \mathbb{E}\left[u\left(W_1\right)\right]. \tag{EU-OP}$$

u^{-1} is increasing for any $u \in \mathcal{U}$, thus an equivalent problem to (EU-OP) is given by

$$\sup_{0 \leq w \leq W_0} u^{-1}\left(\mathbb{E}\left[u\left(W_1\right)\right]\right). \tag{CE-OP}$$

Now, we rather want to analyze this one-period optimal investment model using a criterion based on OEU, namely

$$\sup_{0 \leq w \leq W_0} S^u\left(W_1\right) \tag{OEU-OP}$$

which, due to cash-invariance of OEU, is equivalent to

$$\sup_{0 \leq w \leq W_0} \sup_{\eta \in \mathbb{R}} F(w,\eta) + \beta R_f W_0,$$

with

$$F(w,\eta) = -\beta(R_f w + \eta) + \alpha u^{-1}\left(\mathbb{E}\left[u\left(Rw + \eta\right)\right]\right).$$

When we compare the solution of (OEU-OP) to the solutions of (CE-OP) and of (EU-OP), respectively, we get:

4.26 Proposition. *For any $u \in \mathcal{U}$ we have*

$$\sup_{0 \leq w \leq W_0} S^u(W_1) \geq \alpha \sup_{0 \leq w \leq W_0} u^{-1}\left(\mathbb{E}\left[u\left(W_1\right)\right]\right).$$

Proof. In the present setting it is assured that $W_1 \geq 0$ for any possible choice of w. Therefore we can follow from Proposition 4.13 that

$$S^u(W_1) \geq \alpha u^{-1}\left(\mathbb{E}\left[u(W_1)\right]\right)$$

for all $0 \leq w \leq W_0$. Hence we conclude that

$$\sup_{0\leq w\leq W_0} S^u(W_1) \geq \alpha \sup_{0\leq w\leq W_0} u^{-1}\left(\mathbb{E}\left[u\left(W_1\right)\right]\right). \qquad \square$$

The comparison of the solutions of (OEU-OP) and (EU-OP) is a direct consequence of the previous proposition:

4.27 Corollary. *For any $u \in \mathcal{U}$ such that $u(x) \leq x$ for all $x \in \mathbb{R}_+$ we have*

$$\sup_{0\leq w\leq W_0} S^u(W_1) \geq \alpha \sup_{0\leq w\leq W_0} \mathbb{E}\left[u\left(W_1\right)\right].$$

The following corollary addresses the impact of risk aversion on the solution of (OEU-OP). It is an immediate result of Proposition 4.11.

4.28 Corollary. *If $u,v \in \mathcal{U}$ are such that u is more concave than v then*

$$\sup_{0\leq w\leq W_0} S^u(W_1) \leq \sup_{0\leq w\leq W_0} S^v(W_1).$$

Thus the solution of (OEU-OP) is reduced if the investor switches from a utility function v to a more concave utility function u. Moreover, due to Theorem 4.3, we know that S^u fulfills monotonicity. Consequently, if the safe return R_f increases (which leads to an increase in W_1), then also $\sup_{0\leq w\leq W_0} S^u(W_1)$ increases.

The partial derivatives of $F(w,\eta)$ are (using dominated convergence to justify the interchange of differentiation and expectation):

$$\frac{\partial F}{\partial w} = -\beta R_f + \alpha\frac{\mathbb{E}\left[ru'\left(Rw+\eta\right)\right]}{g(w,\eta)},$$
$$\frac{\partial F}{\partial \eta} = -\beta + \alpha\frac{\mathbb{E}\left[u'\left(Rw+\eta\right)\right]}{g(w,\eta)},$$

with

$$g(w,\eta) = u'\left(u^{-1}\left(\mathbb{E}\left[u\left(Rw+\eta\right)\right]\right)\right).$$

An implicit solution to (OEU-OP) is given by

$$\bar{w}_{\mathrm{OEU}} = \frac{1}{R_1-R_2}\left(\left(u'\right)^{-1}\left(\frac{\beta}{\alpha}\frac{g(\bar{w}_{\mathrm{OEU}},\bar{\eta}_{\mathrm{OEU}})(R_f-R_2)}{(1-p)(R_1-R_2)}\right) - \left(u'\right)^{-1}\left(\frac{\beta}{\alpha}\frac{g(\bar{w}_{\mathrm{OEU}},\bar{\eta}_{\mathrm{OEU}})(R_1-R_f)}{p(R_1-R_2)}\right)\right),$$
$$\bar{\eta}_{\mathrm{OEU}} = -R_1\bar{w}_{\mathrm{OEU}} + \left(u'\right)^{-1}\left(\frac{\beta}{\alpha}\frac{g(\bar{w}_{\mathrm{OEU}},\bar{\eta}_{\mathrm{OEU}})(R_f-R_2)}{(1-p)(R_1-R_2)}\right),$$

thus the optimal amount of money invested in the risky asset can be regarded as

$$w^*_{\mathrm{OEU}} = \begin{cases} 0, & \text{if } \bar{w}_{\mathrm{OEU}} \leq 0, \\ \bar{w}_{\mathrm{OEU}}, & \text{if } 0 < \bar{w}_{\mathrm{OEU}} < W_0, \\ W_0, & \text{if } \bar{w}_{\mathrm{OEU}} \geq W_0 \end{cases}.$$

Let us compare the optimal investment in the risky asset for (OEU-OP) and (EU-OP) for the exponential utility function.

4.29 Example. For $u(x) = \frac{1}{\gamma}(1 - \exp(-\gamma x))$, $\gamma > 0$, $\bar{w}_{\text{OEU}}$ can be explicitly solved:

$$
\begin{aligned}
\bar{w}_{\text{OEU}} &= \frac{1}{R_1 - R_2}\left(-\frac{1}{\gamma}\ln\left(\frac{\beta}{\alpha}\frac{g(\bar{w}_{\text{OEU}},\bar{\eta}_{\text{OEU}})(R_f - R_2)}{(1-p)(R_1 - R_2)}\right) + \frac{1}{\gamma}\ln\left(\frac{\beta}{\alpha}\frac{g(\bar{w}_{\text{OEU}},\bar{\eta}_{\text{OEU}})(R_1 - R_f)}{p(R_1 - R_2)}\right)\right)\\
&= -\frac{1}{\gamma}\frac{1}{R_1 - R_2}\ln\left(\frac{\frac{\beta}{\alpha}\frac{g(\bar{w}_{\text{OEU}},\bar{\eta}_{\text{OEU}})(R_f - R_2)}{(1-p)(R_1 - R_2)}}{\frac{\beta}{\alpha}\frac{g(\bar{w}_{\text{OEU}},\bar{\eta}_{\text{OEU}})(R_1 - R_f)}{p(R_1 - R_2)}}\right)\\
&= -\frac{1}{\gamma}\frac{1}{R_1 - R_2}\ln\left(\frac{p}{1-p}\frac{R_f - R_2}{R_1 - R_f}\right).
\end{aligned}
$$

(EU-OP), on the other hand, takes the form

$$
\begin{aligned}
\sup_{0\le w\le W_0}\mathbb{E}\left[u\left(W_1\right)\right] &= \sup_{0\le w\le W_0}\mathbb{E}\left[\frac{1}{\gamma}(1 - \exp(-\gamma W_1))\right]\\
&= \sup_{0\le w\le W_0}\mathbb{E}\left[\frac{1}{\gamma}(1 - \exp(-\gamma(R_f(W_0 - w) + Rw))\right]\\
&= \sup_{0\le w\le W_0} G(w),
\end{aligned}
$$

where we denote

$$
G(w) := \frac{1}{\gamma}\left(1 - \exp(-\gamma R_f(W_0 - w))\mathbb{E}\left[\exp(-\gamma R w)\right]\right).
$$

Thus

$$
G'(w) = \exp(-\gamma R_f(W_0 - w))\mathbb{E}\left[(R - R_f)\exp(-\gamma R w)\right]
$$

and the first-order condition for an interior optimizer $\bar{w}_{\text{EU}}$ is given by

$$
(1-p)(R_1 - R_f)\exp(-\gamma R_1 \bar{w}_{\text{EU}}) + p(R_2 - R_f)\exp(-\gamma R_2 \bar{w}_{\text{EU}}) = 0.
$$

which is equivalent to

$$
\exp(-\gamma(R_1 - R_2)\bar{w}_{\text{EU}}) = \frac{p(R_f - R_2)}{(1-p)(R_1 - R_f)},
$$

and consequently

$$
\bar{w}_{\text{EU}} = -\frac{1}{\gamma}\frac{1}{R_1 - R_2}\ln\left(\frac{p}{1-p}\frac{R_f - R_2}{R_1 - R_f}\right).
$$

Thus the optimal solutions to (OEU-OP) and (EU-OP) coincide for expected utility investors.

We conclude that the functional S^u suits for setting up an optimal investment problem that is similar to a classical expected utility principle while S^u additionally satisfies economically meaningful properties such as monotonicity, shift additivity and preservation of cash amounts.

CHAPTER 5

Implied risk aversion: An alternative rating system for retail structured products

Structured products for retail clients, or *retail structured products (RSPs)*, are an important asset class for retail investors. Such products, by securitizing the associated risks, open up the possibility to invest in a large variety of national and international markets. The universe of available RSPs is large, and RSPs with widely varying risk profiles, payoff structures and underlying risks are abundant. Therefore, particularly for less sophisticated retail clients, it is important to provide a simple, easily understandable "rating" of RSPs that indicate whether such a product is attractive, by taking into account both the upside potential and the downside risk of the RSP's cash flows. This is highlighted by the findings of Wallmeier [135], who shows that many retail clients are not fully aware of the risks involved in such contracts, and those of Cao and Rieger [27], who demonstrate that simple rules based on expected payoffs and downside risk measures such as value at risk (V@R) may be misleading. This is particularly important since the ratings of *Deutscher Derivate Verband (DDV)*, which is the leading provider of rating information in the German RSP market, are based on V@R. In addition, retail investors may be subject to behavioral biases including, for instance, probability misestimation; see Rieger [120]. Moreover, due to, e.g., information asymmetries, issuers may be able to charge margins or fees that can render some products unattractive; see Henderson and Pearson [85]. However, retails clients are not necessarily able to realize this due to potentially incorrect beliefs, see Hens and Rieger [86], or low informational efficiency, see Schroff, Meyer and Burghof [125]. This overpricing, which is addressed in detail in Section 5.1, is increasing with product complexity, see Celerier and Vallee [29], and might become problematic particularly for retail investors, as individuals with lower competence levels are drawn to such payoffs; see Breuer and Perst [25] for a case study with reverse convertible bonds. However, retail investors are also at a disadvantage for simpler payoffs such as vanilla or barrier options; see Rosetto and van Bommel [122] or Wilkens and Stoimenov [141].

Therefore the objective of this chapter is to provide a comprehensive, integrated risk measure for (net payoffs of) RSPs as an alternative to existing rating methodologies. Building on the notion of optimal expected utility risk measures as introduced in Chapter 4, we introduce implied risk aversion as a coherent RSP rating methodology. In contrast to purely downside risk measures such as V@R or AV@R, implied risk aversion takes into account both the upside potential and the downside risks of such products. In addition, implied risk aversion is easily interpreted in terms of an individual investor's risk aversion: A product is attractive (unattractive) for an investor if its implied risk aversion is higher (lower) than his individual risk aversion. We illustrate our approach in a case study with more than 15,000 warrants on DAX® and find that implied risk aversion is able to identify favorable products; in particular, we show that implied risk aversion is not necessarily increasing with respect to the strikes of call warrants, and that it confirms the intuitively obvious fact that put warrants are generally not attractive instruments for retail investors that assume positive expected returns.

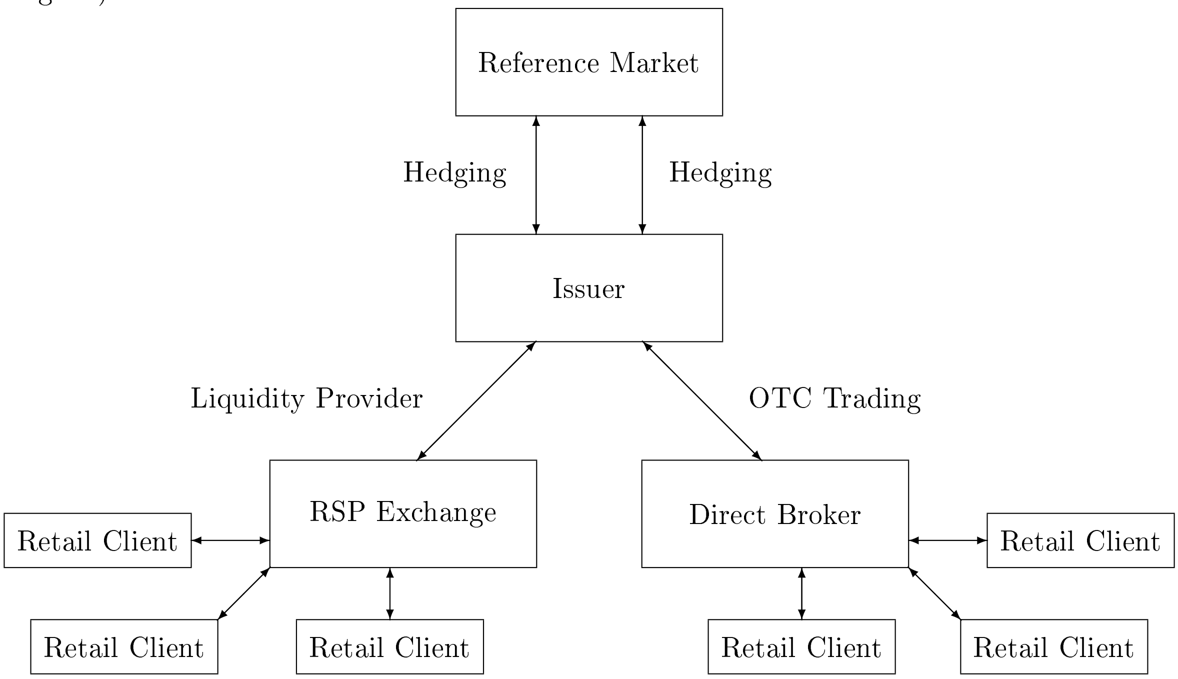
The chapter is structured as follows: In Section 5.1 we provide a brief overview of the German market for RSPs. Section 5.2 is the core theoretical part of this chapter: We first discuss some generalities on ratings and risk measures and introduce the novel notion of fully supported risk measures in Section 5.2.1. In Section 5.2.2 we point out some weaknesses of V@R-based evaluations of RSPs. In Section 5.2.3 we propose implied risk aversion as an alternative rating of RSPs that overcomes many of the drawbacks of existing approaches. We introduce a rating system for RSPs based on implied risk aversion in Section 5.2.4 and relate this system to real-world risk aversion in Section 5.2.5. Section 5.3 presents several models for asset dynamics that we use for our simulations and illustrations in Section 5.4, where we investigate in detail a total of 15,377 warrants on the German blue-chip index DAX®.

This chapter is an extended version of Fink, Geissel, Sass and Seifried [61] which has been submitted to *Journal of Banking and Finance*.

5.1 Structured retail warrants: The German market

Before we address the measurement and quantification of risks in RSPs, we briefly review the institutional background and structure of the German market for RSPs. In this context, a *structured product* is a package of derivatives in the format of a bearer bond (i.e., with counterparty risk) issued by a bank or an investment house. The contract details are usually fixed in a prospectus that forms the legal basis of the corresponding RSP. The first products issued were classical *warrants*, which are basically cash-settled options with the respective issuer's credit risk on top. Over time, more complex types of products have evolved. This development peaked around the Lehman crisis, which can be viewed as a structural break: dwindling demand by retail investors due to fallen confidence in the credit worthiness of banks, combined with ambitious efforts by European and German regulating authorities (ESMA and BaFin) have led to a consolidation of the RSP market. With a total outstanding amount of around EUR 75.2 billion (April 2015, see `http://www.derivateverband.de/ENG/Statistics/MarketVolume`), the market for structured products remains a significant part of Germany's financial market for retail investors. Simpler product types such as classical warrants account for 16.1 percent of the EUR 5 billion total monthly trading volume at RSP exchanges (April

Figure 5.1: Sketch of the German RSP market (client-client trading on exchanges is negligible).



2015, see `http://www.derivateverband.de/ENG/Statistics/StockExchangeTurnover`), and it appears safe to assume an even bigger share in the direct trading/OTC domain.

The issuer of a RSP is typically the sole provider of liquidity in its products; see Figure 5.1. Clients have two ways of buying: either they trade directly/OTC via their respective brokers, or they place an order at an exchange – Frankfurt (Börse Frankfurt) and Stuttgart (EUWAX) are the most important ones. Here, in theory, clients can also trade with each other directly, but the amount of this type of trading is negligible; most of the liquidity is provided by the relevant issuer. Trading hours for direct trade are usually from 08:00 – 22:00 local time, with possible differences for specific product types or underlyings. Exchanges, on the other hand, open at 09:00 and close at 20:00 local time, and issuers have themselves committed to providing a certain amount of liquidity if their products are listed in the premium segments (this is usually the case, especially for DAX® warrants). When a client's order is executed, the issuer will hedge, for instance, via a larger listed-option exchange; for the DAX® investigated in this chapter, this would most likely be EUREX.

The market for RSPs has some unique features and is in itself of special interest for researchers – an excellent overview of the recent literature can be found in Entrop, Schober and Wilkens [58]. As noted above, although there exists some client-to-client trading, the large majority of trades is carried out directly between a client and the issuing investment bank; trading an RSP with a financial institution other than the issuing party is not possible. This opens a potentially huge profit opportunity for these institutions – which is mostly only restricted due to the fact that falling trading volume in recent years have made the market highly competitive; see Döhrer, Johanning, Steiner and Völkle [51] for a current and detailed analysis of priced margin (PnL) with respect to available product types. In particular, it is difficult to set up arbitrage positions using RSPs. This implies that RSPs are typically held for the medium or long term, often until maturity, and highlights the importance of RSP ratings for buyers of these products.

There exists a growing literature on RSPs for retail clients. Early contributions focused on the overpricing mentioned above (i.e., priced-in PnL to be earned by the selling banks), starting with Chen and Kensinger [30] and Chen and Sears [31] for the US market and Wilkens, Erner and Röder [140] for Germany. Later on, Stoimenov and Wilkens [127] put forth a lifecycle hypothesis, i.e., the notion that overpricing diminishes over time when the issuer amortizes his profits by selling the products around their issuance time. A related literature investigates inter- and intraday pricing effects for 'turbos' (basically, barrier options), see [58], as well as classical warrants; see, e.g., ter Horst and Veld [130] for a price comparison between warrants and classical exchange-listed options (in the Netherlands), Schmitz and Weber [124] on retail investor behavior or Baule and Blonski [15] on issuer pricing strategies based on clients' demand.

5.2 Risk classifications of retail structured products

In this section we provide the theoretical framework and background for ratings of RSPs. The main tool is the well-known notion of *risk measures* as presented in Chapter 2. We use an optimal expected utility risk measure as introduced in Chapter 4, and introduce the associated implied risk aversion as an alternative rating system for RSPs.

5.2.1 Risk measures for retail structured products

We refer to Chapter 2 for the formal definitions of a *financial position* and the *risk-free discount factor* β. In particular, in the context of RSP investments, $X(\omega) > 0$ means the investor has made a profit X.

To assess the risk-return profile of a given RSP, we are interested in a map

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

that represents the "riskiness" of its payoff less its costs, X. We consider ρ to be a risk measure, hence the natural conditions for ρ are given in Definition 2.1. Note that positive homogeneity is a crucial property for our purposes: The rating of an RSP should not depend on the amount invested, which means that the risk measure scales linearly in the net payoff.

In the following we recall some features of V@R and AV@R as defined in Example 2.4. While average value at risk is a coherent measure, value at risk fails to be convex. Both share the property that they focus exclusively on the downside of the payoff distribution, which obviously limits their scope for investors that wish to evaluate the risk-return profiles of RSPs. Formally, neither value at risk nor average value at risk are fully supported in the sense of Definition 5.1 below. Therefore we rather turn to risk measures that take into account the entire distribution of the financial position; these are typically utility-based; see Chapter 3 for some well-known examples of utility-based risk measures.

The *optimal expected utility risk measure (OEU)* as introduced in Chapter 4 constitutes another class of utility-based risk measures. OEU is a convex risk measure for any $u \in \mathcal{U}$; see Theorem 4.3. Moreover, if u has constant relative risk aversion (CRRA), i.e.,

$$u(x) = \tfrac{1}{1-\gamma}\left(x^{1-\gamma} - 1\right),$$

for some $\gamma > 0$, $\gamma \neq 1$, then ρ^u is a coherent risk measure; see Theorem 4.16. In this case, the supremum in (OEU) is attained uniquely at $\eta_X^* \in [-X_{\min}, \infty)$ by Proposition 4.7 (ii). Recall that to the best of our knowledge, OEU is the only existing utility-based risk measure that is (non-trivial and) coherent if u has constant relative risk aversion, which is a reasonable assumption; see Section 3.2.

Investments in RSPs can lead to significant losses on the one hand, and to potentially large gains on the other. In particular, the investor may be willing to tolerate a significant downside risk if the product also offers a sufficiently high expected return, see Armbruster and Delage [5]. A fair assessment of a RSP's risk-return profile should therefore take into account both the up- and the downside of the payoff distribution. More generally, we formalize this idea in the following definition.

5.1 Definition. A risk measure ρ is called *fully supported* if

$$\rho(X) > \rho(Y) \quad \text{for all } X, Y \in \mathcal{X} \text{ such that } X \leq Y, X \neq Y.$$

In contrast to purely downside risk measures, such as value at risk or average value at risk, a fully supported risk measure is able to distinguish between financial positions with identical downsides, but different upsides. This is not necessarily the case for standard risk measures: For instance, the value at risk and average value at risk of a vanilla call and a capped call with the same maturities, strikes and underlying are identical if they are not too deep in the money. We next demonstrate that OEU is fully supported.

5.2 Theorem. *If u is a CRRA utility function, the optimal expected utility risk measure ρ^u is fully supported.*

Proof. For $X, Y \in \mathcal{X}$ with $X \leq Y$, $X \neq Y$ we have

$$\begin{aligned}
\rho^u(Y) &= -\sup_{\eta > -Y_{\min}} \left\{-\beta\eta + \alpha u^{-1}\left(\mathbb{E}\left[u\left(Y + \eta\right)\right]\right)\right\} \\
&\leq -\sup_{\eta > -X_{\min}} \left\{-\beta\eta + \alpha u^{-1}\left(\mathbb{E}\left[u\left(Y + \eta\right)\right]\right)\right\} \\
&\leq \beta\eta_X^* - \alpha u^{-1}\left(\mathbb{E}\left[u\left(Y + \eta_X^*\right)\right]\right) \\
&< \beta\eta_X^* - \alpha u^{-1}\left(\mathbb{E}\left[u\left(X + \eta_X^*\right)\right]\right) \leq \rho^u(X),
\end{aligned}$$

where the strict inequality is due to the fact that $\mathbb{E}[u(X + \eta_X^*)] < \mathbb{E}[u(Y + \eta_X^*)]$. □

In summary, we have demonstrated that OEU is a fully supported, coherent risk measure for every CRRA utility function, and hence satisfies all conditions required for an integrated valuation of RSPs.

5.2.2 Weaknesses of rating RSPs with V@R

V@R is a well-established methodology in regulatory risk management. In the context of structured products, the European Union approved new V@R-based regulation on mutual funds called UCITS IV (Undertakings for Collective Investment in Transferable Securities, fourth edition) in 2009, see European Union [59] and a regulation on prudential requirements for credit institutions and investment firms in 2013, see European Union [60]. Since 2005, DDV also provides a V@R-based rating of structured products on the German market, based on a fixed level of $\lambda = 0.01$ and a 10-day holding period; see Döhrer, Johanning, Steiner and Völkle [50].

DDV implements their method in six steps:

1. Simulation of the price of the underlying based on a two-year price history and simulation of market parameters and risk factors (such as price of the underlying, dividends, implied volatility, risk-free interest rates, issuer's interest rate and currency risks)

2. Structuring, i.e., decomposition of products into their specific (option) components

3. Evaluation of the product at the initial date (P_0)

4. Evaluation of the product at the end of the holding period (P_T) based on the simulation in (1.)

5. Calculation of the product's return distribution

6. Derivation of the 1% V@R at the end of the holding period to obtain the price risk of the RSP

In addition, DDV separately computes further risk contributions (currency risk, volatility risk, interest rate risk, counterparty risk) and aggregates these into their V@R assessments; see `http://www.derivateverband.de/ENG/Transparency/ValueAtRisk`. In the following we illustrate some potential drawbacks of such V@R-based rating systems. We consider structured products

(i) where for a given V@R based rating, the largest achievable return is not bounded.

(ii) which have the same V@R but have entirely different payoffs.

(iii) where the V@R leads to a product which is not obviously the better choice.

5.3 Example. (i) Cao and Rieger [27] construct RSPs from four call options that satisfy a given V@R constraint, while their expected returns can reach any size. Thus an investor that naively selects the RSP with the the largest expected return in a given risk class might opt for a product that is not truly appropriate for his risk preferences, overlooking potentially high risks.

(ii) This scenario is fulfilled by (almost) any investment of, say EUR 10,000, in short-term, at-the-money call warrants with payoff X_i^{call} on the same underlying. Obviously, if we consider V@R at maturity, we get

$$V@R^{0.01}(X_i^{\text{call}}) = 10{,}000 \text{ for all } i.$$

For example, consider the following calls on DAX®:

Table 5.1: 5-day call warrants on DAX® rated equally.

WKN	Issuer	Strike	Ask	Underlying	$V@R$
XM2HA5	Deutsche Bank	11,300	5.40	11,830.66	10,000
XM2HAK	Deutsche Bank	12,000	0.32	11,831.29	10,000

Clearly, the payoffs of these calls are qualitatively different. For instance, the likelihood of total loss is higher the higher the strike. On the other hand, warrants with higher strikes offer more leverage and may therefore be attractive for investors that value the RSP's upside potential. In summary, it is quite clear that in such a situation rating RSPs with V@R does not give investors helpful information for their investment decisions.

(iii) Consider short-term call warrants that are deep in the money. Since value at risk, by definition, only takes account of the worst 1 percent of the simulated outcomes of the underlying, we obtain a smaller V@R the smaller the strike of the call option. However, one might argue that the payout profiles of these products are nearly identical, and that the worst-case scenario of a total loss only applies in case of a severe market crash, which will affect all warrants equally. In additional, higher strike prices lead to increased leverage, which might make these products more attractive to some investors. The following pair of call warrants on DAX® illustrates this point:

Table 5.2: 5-day call warrants on DAX® with potentially misleading V@Rs.

WKN	Issuer	Strike	Ask	Underlying	$V@R$
XM2H9D	Deutsche Bank	10,000	18.32	11,830.38	3,339
XM2H9V	Deutsche Bank	10,800	14.32	11,830.46	5,927

The above examples illustrate that rating systems may mislead unsophisticated investors who are not able to take into account the exact payoff profile or the complete return distribution. However, it is exactly for such investors that RSPs are designed. Thus there appears to be a need for an alternative rating system that is suitable for unsophisticated retail clients.

Finally we give an example of a simple Bernoulli-type payoff in order to illustrate that, in contrast to V@R and AV@R, OEU is sensitive with respect to changes in the amount of both, potential losses and gains of X.

5.4 Example. We consider the financial position

$$X = \begin{cases} -l & \text{with } p = 0.5, \\ g & \text{with } p = 0.5, \end{cases}$$

where $l \geq 0$ is the potential loss and $g \geq 0$ is the potential gain of the position. For this example we choose $\alpha = 0.9$, $\beta = 0.95$, $\lambda = 0.05$ and calculate OEU for a CRRA utility function with risk aversion parameter $\gamma = 3$. Figures 5.2, 5.3 and 5.4 illustrate OEU, V@R and AV@R as functions of l and g.

Figure 5.2: V@R as a function of l and g.

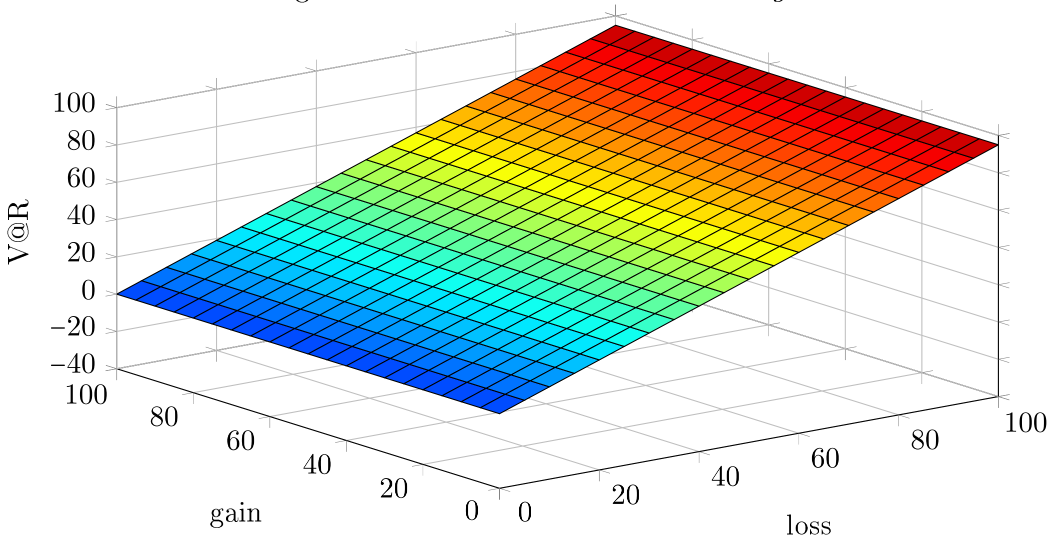


Figure 5.3: AV@R as a function of l and g.

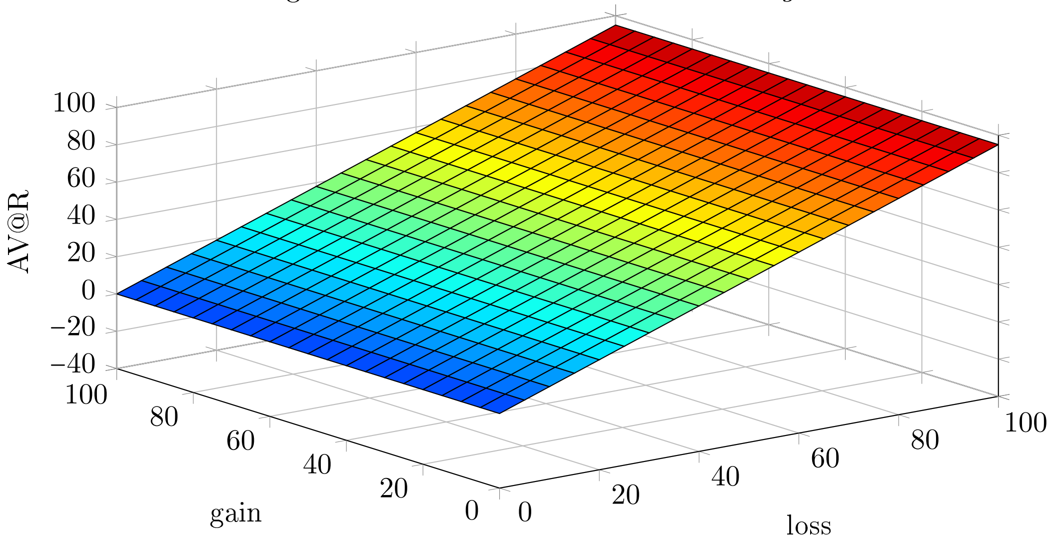
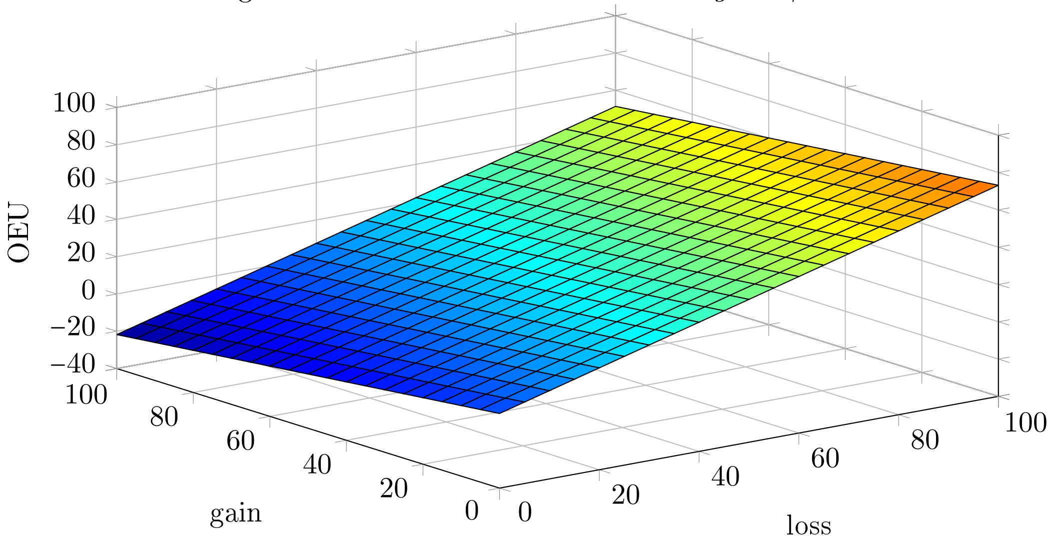


Figure 5.4: OEU as a function of l and g for $\gamma = 3$.



Figures 5.2 and 5.3 illustrate that V@R and AV@R do not offer any qualitative information to an investor as they just equal the potential loss of X. By contrast, we note that OEU increases for decreasing amounts of potential gains as well as for ever larger potential losses; see Figure 5.4. In particular, ρ^u is negative if the investor values the upside potential more than possible losses; a negative risk assessment seems obviously reasonable if, for example, $l = 0$.

5.2.3 Implied risk aversion

In the following we develop a utility-based rating system for RSPs. As demonstrated in Section 5.2.1, OEU is a fully supported and coherent risk measure for a CRRA utility function with relative risk aversion $\gamma > 0$. Thus we base our rating system on OEU of the form

$$\rho^u(X,\gamma) = \rho^u(X) = -\max_{\eta > -X_{\min}} \left\{ -\beta\eta + \alpha\mathbb{E}\left[(X+\eta)^{1-\gamma}\right]^{\frac{1}{1-\gamma}} \right\}, \quad X \in \mathcal{X}.$$

As explained in a short note below Example 4.10, we set $u(x) = \ln(x)$ for $\gamma = 1$, so

$$\rho^u(X,1) = -\max_{\eta > -X_{\min}} \left\{ -\beta\eta + \alpha \exp\left(\mathbb{E}\left[\ln(X+\eta)\right]\right) \right\}, \quad X \in \mathcal{X}.$$

By Jensen's inequality, $\rho^u(X,\gamma)$ is strictly increasing in γ unless X is deterministic; for deterministic X, $\rho^u(X,\gamma)$ is obviously constant. Hence we can rate the product X by the largest value $\gamma > 0$ such that an investor with relative risk aversion γ considers X attractive:

5.5 Definition. For any $X \in \mathcal{X}$ we define the *implied risk aversion* $\gamma_0(X)$ via

$$\gamma_0(X) := \inf\{\gamma > 0 : \rho^u(X,\gamma) \geq 0\}$$

where by convention $\inf \varnothing := \infty$.

Note that, if $\rho^u(X,\gamma) = 0$ for some $\gamma > 0$, then that value is unique and $\gamma_0(X) = \gamma$. On the other hand, we have $\gamma_0(X) = \infty$ if $\rho^u(X,\gamma) < 0$ for all $\gamma > 0$, and $\gamma_0(X) = 0$ if $\rho^u(X,\gamma) \geq 0$ for all $\gamma > 0$. Intuitively, an investor with risk aversion $\gamma_0(X)$ is indifferent between the financial position X and a zero investment; the position X is favorable (unfavorable) for any investor with $\gamma < \gamma_0(X)$ (respectively, $\gamma > \gamma_0(X)$). In particular, note that the smaller $\gamma_0(X)$, the riskier (i.e., less attractive) the product. Here and in all that follows, we fix the remaining parameters

$$\alpha = \frac{1}{1 + 0.2 \cdot T/365} \quad \text{and} \quad \beta = \frac{1}{1 + 0.005 \cdot T/365}.$$

This corresponds to a risk-free rate of 0.5% p.a. and a subjective rate of time preference of 20% p.a.,[1] where T denotes the investor's time horizon in days (typically, $T = 90$ days).

5.6 Remark. Alternatively, one might analogously construct a rating based on utility functions with constant absolute, rather than relative, risk aversion or, in fact, any parametric family of utility functions. However, as shown in Section 5.2.1, implied risk aversion as defined above is the only choice that leads to a coherent risk measure, and thus makes the implied rating independent of the amount invested.[2]

We illustrate implied risk aversion with a simple RSP example.

5.7 Example. Consider a hypothetical investment of EUR 10,000 into a call warrant on DAX®:

Table 5.3: 5-day call warrant.

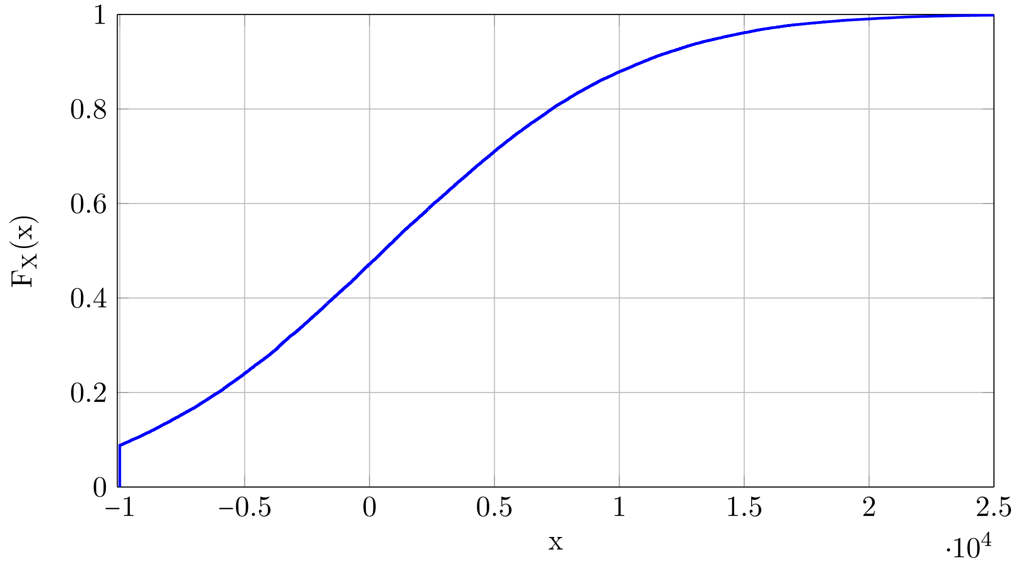
WKN	Issuer	Strike	Ask	Underlying
XM2HA9	Deutsche Bank	11,500	3.50	11,829.89

The price of the underlying is simulated as a geometric Brownian motion with drift and volatility parameters estimated on the basis of historical data; see Section 5.3 for details. The cumulative distribution of the product's net return looks as follows:

[1] We show in Section 5.4.1 below that our results are robust with respect to alternative specifications of α.

[2] Even more generally, the approach here can also be used for parametric families of downside risk measures; e.g., one may define $\lambda_0^{\mathrm{V@R}}(X) := P(X \geq 0)$ as the probability of not making a loss from an investment X, or $\lambda_0^{\mathrm{AV@R}}$ as the level such that the mean V@R for all $\lambda \geq \lambda_0^{\mathrm{AV@R}}$ equals zero. Such constructions suffer from the fact that downside risk measures are not fully supported.

Figure 5.5: Cumulative distribution function of X.

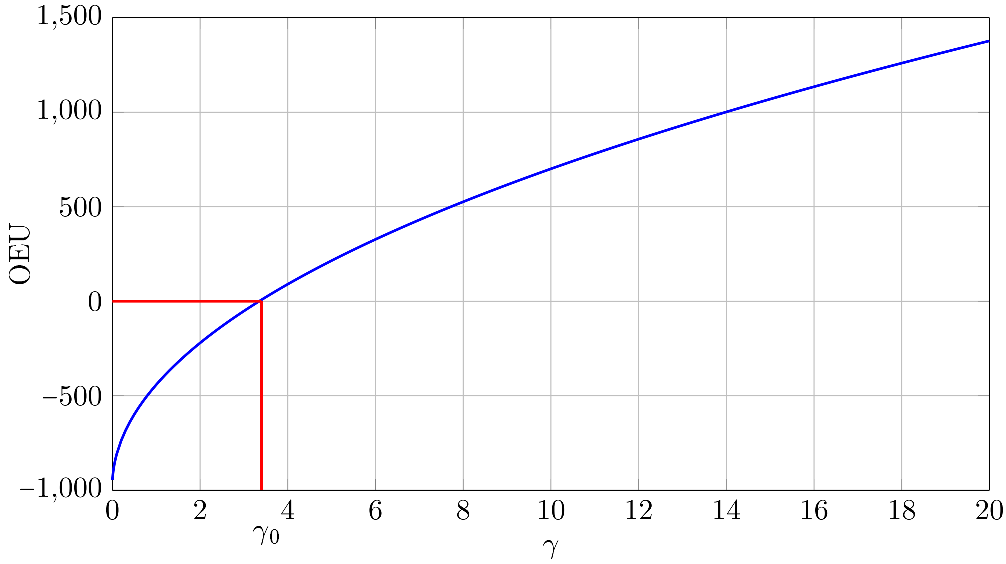


Clearly, an evaluation of the product with V@R yields the trivial result

$$V@R^{0.01}(X) = 10{,}000.$$

By contrast, let us consider implied risk aversion: Figure 5.6 displays $\rho^u(X,\gamma)$ as a function of the investor's relative risk aversion γ:

Figure 5.6: OEU depending on γ.



In particular, in this example the implied risk aversion of X is given by

$$\gamma_0(X) = 3.1703.$$

The following result summarizes the main properties of implied risk aversion and provides some justification to use implied risk aversion for the classification of RSPs.

5.8 Theorem. *For all $X, Y \in \mathcal{X}$ the following properties hold:*

(i) *γ_0 fulfills monotonicity: $\gamma_0(X) \leq \gamma_0(Y)$ whenever $X \leq Y$.*

(ii) *γ_0 supports diversification: $\gamma_0(\lambda X + (1-\lambda)Y) \geq \min\{\gamma_0(X), \gamma_0(Y)\}$ for every $0 \leq \lambda \leq 1$.*

(iii) *γ_0 is invariant with respect to the amount invested: $\gamma_0(\lambda X) = \gamma_0(X)$ for all $\lambda \geq 0$.*

(iv) *γ_0 preserves second-order stochastic dominance: $\gamma_0(X) \leq \gamma_0(Y)$ if $\mathbb{E}[v(X)] \leq \mathbb{E}[v(Y)]$ for all increasing, concave functions $v : \mathbb{R} \to [-\infty, \infty)$.*

Proof. The stated properties follow from the definition of implied risk aversion and some results on OEU from Chapter 4:

(i) Let $X, Y \in \mathcal{X}$ such that $X \leq Y$. Due to Theorem 4.3, for any $\gamma > 0$, we have

$$\rho^u(X, \gamma) \geq \rho^u(Y, \gamma). \tag{5.1}$$

Monotonicity of γ_0 can easily be checked for $X, Y \in \mathcal{X}$ such that $\gamma_0(X), \gamma_0(Y) \in \{0, \infty\}$ by Definition 5.5. In order to prove monotonicity for $0 < \gamma_0(X), \gamma_0(Y) < \infty$, note that (5.1) particularly implies

$$\rho^u(X, \gamma_0(Y)) \geq \rho^u(Y, \gamma_0(Y)) \geq 0,$$

where the last inequality follows from Definition 5.5. Consequently,

$$\gamma_0(Y) \in \{\gamma > 0 : \rho^u(X, \gamma) \geq 0\}$$

and thus $\gamma_0(X) \leq \gamma_0(Y)$ due to Definition 5.5.

(ii) Without loss of generality we assume that $\gamma_0(X) = \min\{\gamma_0(X), \gamma_0(Y)\}$. If $\gamma_0(X) = 0$ the conclusion follows easily. We therefore only consider the case $0 < \gamma_0(X)$ and deny the conclusion and suppose that $\gamma_0(\lambda X + (1-\lambda)Y) < \gamma_0(X)$. Then, by Definition 5.5:

$$\rho^u(X, \gamma_0(\lambda X + (1-\lambda)Y)) < 0 \quad \text{and} \quad \rho^u(Y, \gamma_0(\lambda X + (1-\lambda)Y)) < 0,$$

thus, due to Theorem 4.3:

$$\begin{aligned}\rho^u(\lambda X + (1-\lambda)Y, \gamma_0(\lambda X + (1-\lambda)Y)) &\leq \lambda \rho^u(X, \gamma_0(\lambda X + (1-\lambda)Y)) \\ &\quad + (1-\lambda)\rho^u(Y, \gamma_0(\lambda X + (1-\lambda)Y)) \\ &< 0\end{aligned}$$

which contradicts the definition of $\gamma_0(\lambda X + (1-\lambda)Y)$ and thereby concludes the proof.

(iii) From Theorem 4.16 we know that for any $\gamma > 0$

$$\rho^u(\lambda X, \gamma) = \lambda \rho^u(X, \gamma) \quad \text{for all } \lambda \geq 0.$$

Therefore

$$\inf\{\gamma > 0 : \rho^u(X, \gamma) \geq 0\} = \inf\{\gamma > 0 : \rho^u(\lambda X, \gamma) \geq 0\} \quad \text{for all } \lambda \geq 0.$$

(iv) Theorem 4.15 implies that for any $\gamma > 0$ we have $\rho^u(X,\gamma) \geq \rho^u(Y,\gamma)$ if $\mathbb{E}[v(X)] \leq \mathbb{E}[v(Y)]$ for all increasing, concave functions $v : \mathbb{R} \to [-\infty,\infty)$. Thus, analogously to the proof of monotonicity, we conclude that

$$\gamma_0(X) \leq \gamma_0(Y). \qquad \square$$

Properties (i) and (iv) ensure that implied risk aversion is consistent with "obvious" rankings by monotonicity, while (ii) shows that implied risk aversion encourages diversification to reduce total risk. Importantly, property (iii) guarantees that a product's implied risk aversion is independent of the amount invested into that product; equivalently, the rating of a given product X is the same as that of a positive multiple λX. Thus implied risk aversion assesses only the risk-return tradeoff, while the decision how much capital to invest is left to the investor.

5.2.4 Implied risk aversion as a rating system

DDV categorizes RSPs into one of five risk classes. These are distinguished by their simulated V@R for a hypothetical investment of EUR 10,000 in the relevant RSP; see Table 5.4.

Table 5.4: DDV classification from [50, Figure 5].

risk class	thresholds in %	investor type
1	$0 < V@R^{0.01} \leq 2.5$	risk averse
2	$2.5 < V@R^{0.01} \leq 7.5$	limited willingness to take risks
3	$7.5 < V@R^{0.01} \leq 12.5$	willing to take risks
4	$12.5 < V@R^{0.01} \leq 17.5$	increased willingness to take risks
5	$17.5 < V@R^{0.01} \leq 100$	speculative

We now convert these risk classes into a system based on implied risk aversion. For this purpose, we consider an investment of EUR 10,000 for 90 days in a financial product with a log-normal distribution. The associated drift parameter μ is taken from a maximum likelihood estimation based on a 2-year DAX® price history; the associated volatility σ is computed in such a way that the payoff attains the threshold V@R in Table 5.4. Then we use this log-normal product to determine the implied threshold risk aversion γ_0. The resulting rating scheme is presented in Table 5.5.

Table 5.5: Rating system based on implied risk aversion.

classification	threshold risk aversion	investor type
1	$29.5 < \gamma_0$	risk averse
2	$11.9 < \gamma_0 \leq 29.5$	limited willingness to take risks
3	$6.4 < \gamma_0 \leq 11.9$	willing to take risks
4	$4.1 < \gamma_0 \leq 6.4$	increased willingness to take risks
5	$\gamma_0 \leq 4.1$	speculative

A survey among retail investors in Germany carried out by DDV [47] shows that typical investors' risk preferences cover all five rating classes; see Table 5.6.

Table 5.6: Risk classes among private investors due to [47].

risk class	1	2	3	4	5
share	10.3%	21.9%	24.1%	15.0%	28.7%

We end this section with a brief and incomplete review of the empirical literature on individuals' risk aversion. A large part of the literature, including, among many others, Barsky, Kimball, Juster and Shapiro [14] and Kimball, Sahm and Shapiro [94] is based on laboratory experiments or hypothetical survey questions; field research in this area has been carried out by, for instance, Binswanger [20]. Meyer and Meyer [107], to whom we also refer for an overview, perform a meta-analysis of past studies on the real-world risk aversion and point out that risk aversion for consumption may be up to five times higher than risk aversion for wealth. They therefore adjust previously reported estimates and find relative risk aversion parameters for wealth between 1 and 7, with 3 as a viable "typical" value; see [107, Table 2]. This is in line with other recent studies such as Paravisini, Rappoport and Ravina [110], who find an average relative risk aversion of 2.85 or [94] in data from person-to-person lending platforms, who obtain a value of RRA of 1.64 (when multiplied by the factor 0.2 in the sense of [107]). More extreme values of risk aversion have also been found by, among others, [110], see Table 5.7, and by Janecek [87], who discusses that professional gamblers' risk aversions are in excess of 20, with even higher values for less experienced individuals.

Table 5.7: Distribution of RRA from [110].

fractiles	1	10	25	50	75	90	99
risk aversion	-0.16	0.28	0.56	1.62	3.66	7.29	17.18

5.2.5 Real-world risk aversion

Assume that an investor in RSPs only uses personal wealth which is (given a certain investment horizon) not necessary for his daily spending in any other area of life, but is exclusively intended to be used for investments. We shall call this amount of money *investment wealth* and denote it by W. All of the investment decisions for W are assumed to be made under the individual constant relative risk aversion γ. Thus it is of great importance for the investor to know and follow his personal relative risk aversion in order to make reasonable decisions on offered investment opportunities.

We suggest a self-test for the personal relative risk aversion: Assume that your investment wealth W is EUR 10,000 and that your investment horizon is 90 days, then answer the following question:

What is the maximum amount you are willing to pay for a financial position X with the following payoff:

$$X = \begin{cases} 10{,}000, & \text{with } p_1 = 0.8, \\ 0, & \text{with } p_2 = 0.2 \end{cases} \quad ?$$

Obviously, if you are willing to risk x for the chance of winning 10,000 then x is your *indifference value* of this gamble. We can map x against the degree of relative risk aversion γ by means of the OEU of X with a CRRA utility function with relative risk aversion γ as in Section 5.2.3:

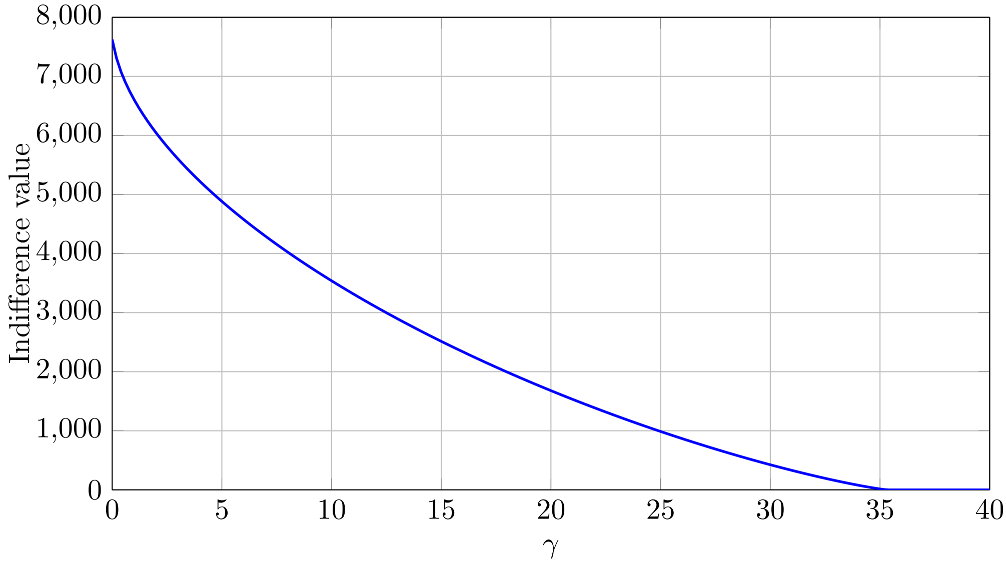
$$\rho^u(X,\gamma) = x.$$

We have given justification for the choice of CRRA utility function in Section 3.2. This assumption can also be understood in the following sense: While an investor may not possess exactly a risk behavior as modeled by the power utility function, we are trying to find the constant relative risk aversion parameter which would best approximate it. Note that coherency of ρ^u is a crucial feature when considering investment decisions within the range of one's investment wealth: The answer to offered gambles of the above type should always be the same proportion of the potential win. To find your personal relative risk aversion, some exemplary conversions and a graphical representation of the relation of x and γ are illustrated in Table 5.8 and Figure 5.7.

Table 5.8: RRA self-test.

x	γ	x	γ	x	γ	x	γ
7,624	0	3,775	9	1,994	18	747	27
6,987	0.5	3,656	9.5	1,913	18.5	690	27.5
6,606	1	3,539	10	1,834	19	634	28
6,304	1.5	3,426	10.5	1,756	19.5	580	28.5
6,045	2	3,315	11	1,679	20	527	29
5,813	2.5	3,207	11.5	1,604	20.5	475	29.5
5,602	3	3,101	12	1,530	21	425	30
5,406	3.5	2,998	12.5	1,458	21.5	376	30.5
5,222	4	2,897	13	1,387	22	328	31
5,048	4.5	2,799	13.5	1,317	22.5	282	31.5
4,884	5	2,702	14	1,249	23	238	32
4,726	5.5	2,607	14.5	1,182	23.5	195	32.5
4,575	6	2,514	15	1,116	24	154	33
4,431	6.5	2,423	15.5	1,051	24.5	115	33.5
4,291	7	2,334	16	988	25	79	34
4,156	7.5	2,247	16.5	926	25.5	45	34.5
4,025	8	2,161	17	865	26	15	35
3,898	8.5	2,077	17.5	805	26.5	0	35.5

Figure 5.7: RRA self-test.



5.3 A simulation for warrants on DAX®

In this section we set up a simulation framework to compute the implied risk aversions for a specific class of RSPs based on the German blue-chip index DAX®. We implement 5 alternative dynamic models for the evolution of this underlying and calibrate each to real-world market data. In Section 5.4 below we apply these models to provide ratings of more than 15,000 RSPs based on DAX® to illustrate our implied risk aversion approach. Before we address RSPs, however, we discuss in detail the models and our estimation and calibration methodologies in this section.

To demonstrate the applicability and robustness of our implied risk aversion approach, we have selected 5 distinct models for the dynamics of DAX®. In addition to a standard Gaussian model, which is also used by DDV in their ratings, we also use more sophisticated models that reflect the so-called "stylized facts" of financial returns time series such as heavy tails, skewness and stochastic volatility; see Cont [39] or Rachev [115] for an overview. Specifically, we use a variance gamma, a normal tempered, an ARMA-GARCH and a historical simulation approach.

- **Model 1: Black-Scholes**
  In the well-known *Black-Scholes model*, see Black and Scholes [21] and Merton [106], the asset price process is a geometric Brownian motion, i.e., satisfies the stochastic differential equation

  $$dS(t) = \mu^* S(t)dt + \sigma^* S(t)dW(t), \quad S(0) \geq 0, \quad t \geq 0,$$

  with W a Brownian motion, $\mu^* \in \mathbb{R}$ and $\sigma^* \geq 0$. This implies that logarithmic returns follow a normal distribution with expectation μ^* and variance $(\sigma^*)^2$, so parameter estimation via *maximum likelihood (MLE)* and simulation of asset paths is straightforward.

- **Model 2: Pure jump variance gamma**
  The *variance gamma (VG) process* is among the most commonly used pure jump processes in finance and was introduced in various versions by Madan, Carr and Chang [101], Madan and Seneta [103], and Madan and Milne [102]. We follow [101] and define a variance gamma process by

  $$VG(t) = \theta^* \Gamma_{\nu^*}(t) + \sigma^* W(\Gamma_{\nu^*}(t)), \quad t \geq 0,$$

  where $\theta^* \in \mathbb{R}$ and Γ_{ν^*} is a gamma process with mean 1 and variance ν^*. The stock price dynamics are given by

  $$S(t) = S(0) \exp(\mu^* t + VG(t) + \omega^* t), \quad S(0) \geq 0, \quad t \geq 0,$$

  where $\mu^* \in \mathbb{R}$ and $\omega^* = {\nu^*}^{-1} \ln(1 - \theta^* \nu^* - 0.5(\sigma^*)^2 \nu^*)$. Since the characteristic function of the variance gamma distribution is known in closed form, the return density required for the parameter estimation via MLE can be obtained by classical Fourier inversion (alternatively, with the closed-form formula of Madan, Carr and Chang [101]). For simulation purposes, a classical Euler-Maruyama scheme can be applied using gamma and normal random variables.

- **Model 3: Normal tempered stable**
  Similarly as in the variance gamma approach, the *normal tempered stable (NTS) model* is based on a time-changed Brownian motion. Instead of a gamma process, a tempered stable subordinator (which is basically special case of an CGMY process, see Carr, Geman, Madan and Yor [28]) is used, which allows for increased flexibility in modeling tails and skewness. NTS processes were proposed by Barndorff-Nielsen and Levendorskii [11] and have been used successfully in various financial applications; see Barndorff-Nielsen and Shephard [12], Kim, Giacometti, Rachev, Fabozzi and Mignacca [90] and Kim and Volkmann [91] among others. In the context of RSPs, NTS processes have been employed by Fink and Mittnik [62] to price quanto warrants and barrier options on the Nikkei 225. Our NTS setup can be described as follows: Let $\mathcal{T} = \{\mathcal{T}(t)\}_{t \geq 0}$ be a tempered stable subordinator, i.e., an a.s. increasing Lévy process with characteristic function

  $$\mathbb{E}\Big[\exp(iv\mathcal{T}(t))\Big] = \exp\left(-\frac{2t(\theta^*)^{1-\alpha^*/2}}{\alpha^*}\Big[(\theta^* - iv)^{\alpha^*/2} - (\theta^*)^{\alpha^*/2}\Big]\right), \quad v \in \mathbb{R}, \quad t \geq 0,$$

  where $\alpha^* \in (0,2)$ is the tail parameter and $\theta^* > 0$ controls the tempering. Then the NTS process is defined via

  $$X(t) = \mu^* t + \beta^*(\mathcal{T}(t) - t) + \sigma^* W(\mathcal{T}(t)), \quad t \geq 0,$$

  with $\mu^*, \beta^* \in \mathbb{R}$ and $\sigma^* \geq 0$. The parameter β^* then controls the skewness. The asset price is given by

  $$S(t) = S(0) \exp(X(t)), \quad S(0) \geq 0, \quad t \geq 0.$$

As in the variance gamma framework, the characteristic function of X is known in closed form, so MLE can be carried out via Fourier inversion. To simulate paths, we follow the approach of Baeumer and Meerschaert [10] and Kawai and Masuda [89]; also see Fink and Mittnik [62]. To obtain realistic stock price paths, we restrict ourselves to $\alpha^* \in [1,2]$ as in [62].

- **Model 4: ARMA-GARCH**
  From the variety of stochastic volatility extensions of the classical Black-Scholes or Lévy setups, we select a classical discrete-time *ARMA-GARCH process*; see Bollerslev [22], Box, Jenkins and Reinsel [24], Engle [57] and Whittle [139]. This is arguably the most well-known and most widely used time series model for financial returns. To stay in line with the Lévy-based models above, we restrict ourselves to an ARMA(1,1)-GARCH(1,1) setup. This implies a total of 6 parameters in addition to those of the error distribution, and log returns $r(t)$ are given by

$$\begin{aligned} r(t) &= a_0^* + a_1^* r(t-1) + \eta(t) + b_1^* \eta(t-1), \\ \eta(t) &= \sigma^*(t)\varepsilon(t), \\ \sigma^2(t) &= \alpha_0^* + \alpha_1^* r(t-1)^2 + \beta_1^* \sigma(t-1)^2, \quad t \in \mathbb{Z}, \end{aligned}$$

  with $\alpha_0^* > 0$, $\alpha_1^* + \beta_1^* < 1$, $|a_1^*| < 1$ and $a_0^*, \beta_1^* \in \mathbb{R}$. The $(\varepsilon(t))_{t\in\mathbb{Z}}$ are i.i.d. and follow a normal, $t-$ or Hansen's skewed-t distribution; see Hansen [82]. Estimations and simulations are carried out using Kevin Sheppard's MFE Toolbox for MATLAB.

- **Model 5: Empirical distribution**
  As a final benchmark, we use a classical historical simulation approach, i.e., we use actual, historical daily returns on DAX®.

To calibrate the underlying parameters of these models, we follow DDV's approach and use daily DAX® returns from the past two years. Our parameter estimates are reported in Tables 5.9 and 5.10.[3] Estimated densities and the QQ-plot for Model 4 are presented in Figures 5.8, 5.9 and 5.10.

Table 5.9: Parameter estimates for Models 1-3.

Parameters	Model 1	Model 2	Model 3
μ^*	0.0007	0.0007	0.0007
σ^*	0.0105	0.0107	0.0106
ν^*	-	0.7322	-
θ^*	-	-0.0010	0.6500
α^*	-	-	1.0000
β^*	-	-	-0.0018

[3] We indicate these parameters with a '*' to avoid confusion with α and β from Section 5.2.3. For Model 4, *Hansen's skewed-t distribution* (with skew parameter 7.0809 and tail parameter -0.0962) turned out to be the best for the error terms based on an evaluation of AIC, BIC, log likelihood and a QQ-plot.

Table 5.10: Parameter estimates for Model 4.

a_0^*	a_1^*	b_1^*	α_0^*	α_1^*	β_1^*
0.0005	0.2075	-0.2494	3.0229e-06	0.1211	0.8620

Figure 5.8: Empirical density of daily DAX® returns vs. densities of Models 1-3.

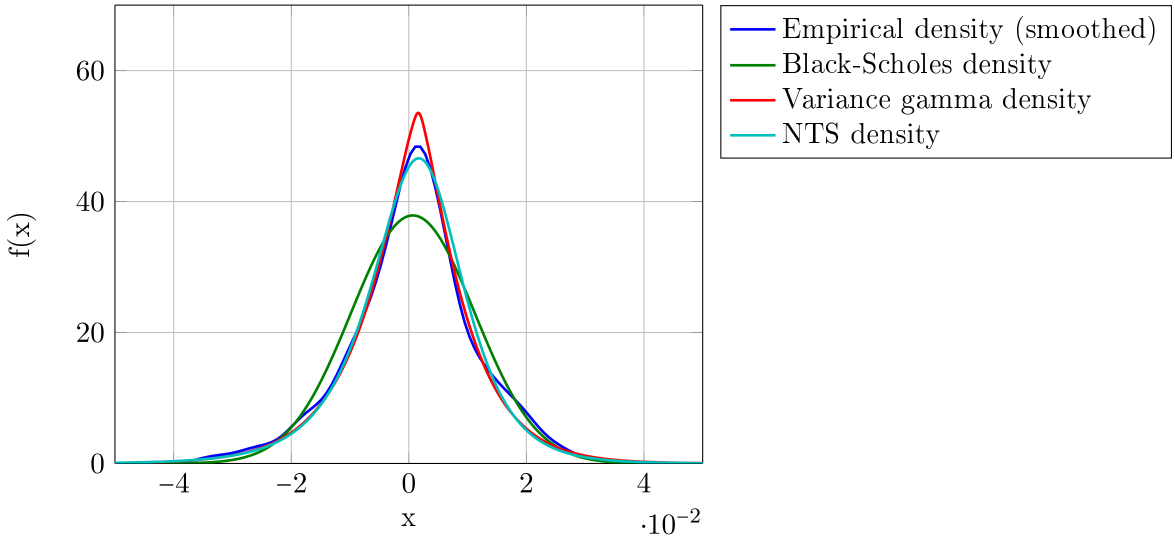


Figure 5.9: Associated densities in the left tails.

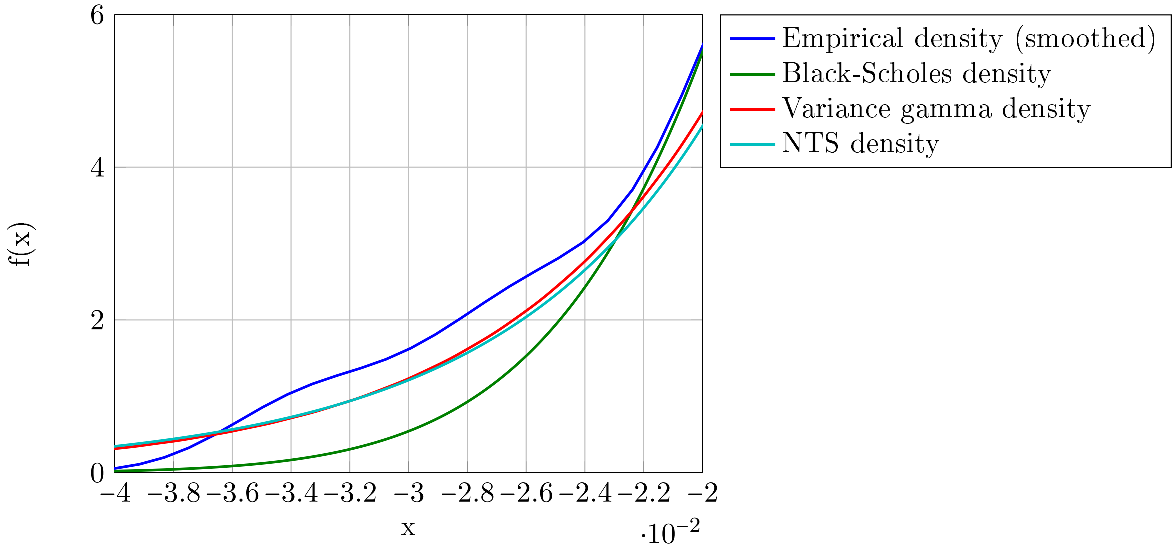
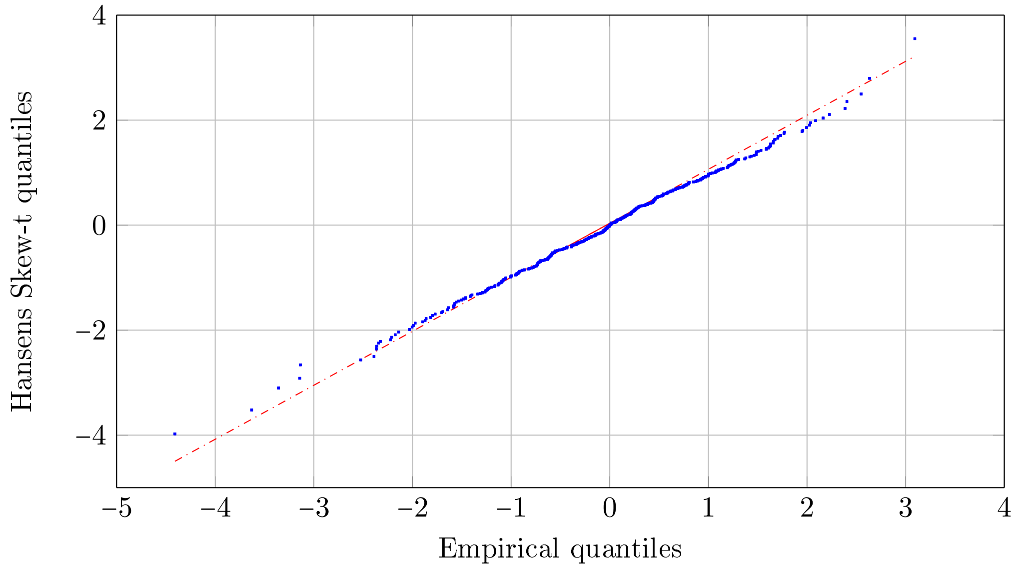


Figure 5.10: QQ-Plot for Model 4 using Hansen's skewed-t distribution with skew parameter 7.0809 and tail parameter -0.0962.



5.4 Case study: Ratings of DAX® warrants

In this section we conduct a case study of our rating system for a specific class of RSPs. We focus on warrants with maturities up to 3 months on DAX® on a given spot date, which form a universe of 15,377 products; an overview of the various issuers/market makers can be found in Table 5.11. In this setting, we illustrate the implied risk aversion rating as proposed in this chapter, and contrast it with the existing V@R approach as used by, e.g., DDV, see Section 5.2.2. To avoid nested simulations and questions of pricing, we focus on holding periods that coincide with the maturity of the respective warrant.[4] All risk measures and ratings are based on the framework outlined in Section 5.2 and are evaluated using simulations with 20,000 paths,[5] with an initial investment of EUR 10,000.

Market prices of RSPs are taken from OnVista. Ask prices and time stamps are extracted from its warrant selection tool on 22 May 2015, a typical trading day, between 16:45 and 17:20 Frankfurt time, resulting in 9,033 call and 9,511 put warrants. After removing errors and outdated quotations, we were left with a total of 15,377 warrants (7,777 calls and 7,600 puts). Based on the individual time stamp of each RSP, we selected the corresponding DAX® prints using tick data from Börse Frankfurt; the spot price is approximately 11,830. In addition, to reduce complexity, maturities were rounded to full days, and we do not distinguish between products that expire on the mid-day or end-of-day auction. In Figure 5.11 we display the strikes and maturities of all warrants in our

[4] This implies in particular that, in DDV's approach, risk factors other than the underlying's market risk are not relevant.

[5] For numerical stability, γ_0-values below 0.0001 are set to zero.

Table 5.11: Overview: RSPs per issuer.

Issuer	Call warrants	Put warrants	Total
BNP Paribas	452	450	902
Citigroup	651	749	1,400
Commerzbank	1,105	1,062	2,167
Deutsche Bank	1,183	985	2,168
DZ Bank	622	456	1,078
Goldman Sachs	447	445	892
HSBC	290	316	606
HypoVereinsbank	930	985	1,915
Interactive Broker	344	356	700
Lang & Schwarz	10	19	29
Société Générale	65	81	146
UBS	942	925	1,867
Vontobel	736	771	1,507

study. It is apparent that the largest share is made up of at-the-money warrants with short maturities. As discussed below, it is exactly for these warrants that implied risk aversion provides a more sensible rating than V@R.

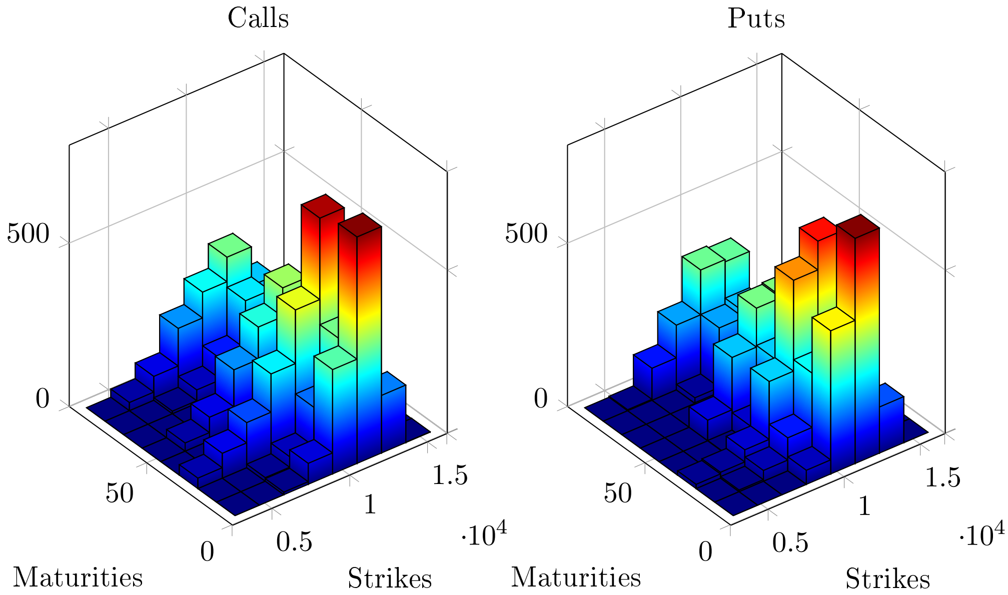
The analysis of this section consists of two parts: First, we investigate the big picture and discuss the resulting ratings and risk classifications of all warrants in our sample. Second, we focus on some hand-picked products to exemplify the potential benefits of using implied risk aversion for RSP ratings.

5.4.1 The bigger picture: Implied risk aversion vs. V@R

In the following we consider call and put warrants separately. Beginning with the analysis of calls, Figure 5.12 displays, for each financial market model from Section 5.3, the distributions of the implied risk aversion parameters (blue, line plot) and the V@Rs (green, bar plot) of all warrants in our sample. To improve visualization, we display a smoothed density for implied risk aversion; the red line represents the threshold between risk classes 4 and 5. Both implied risk aversion and V@R classify the majority of products into the highest risk class. While V@R assigns the maximum value EUR 10,000 to more than 70% of products and groups virtually all into risk class 5, implied risk aversion displays a more nuanced picture: The distribution peaks around the threshold at relative risk aversion between 3 and 4 and exhibits only a small hump at 0. The results are reasonably robust across financial market models; a slight exception is the ARMA-GARCH setup, where we observe a more positive skew in the distribution of γ_0, and a greater number of products with V@R equal to 10,000. We trace these effects to the fact that Hansen's skewed t-distribution is used as an error distribution.

While both rating approaches group the majority of products into the highest risk class, implied relative risk aversion clearly assesses RSP risk in a different way than V@R, and is less granular in doing so. To illustrate, we display the associated risk classifications in Table 5.12. In particular, while there are no products in V@R classes 1, 2 or 3, implied

Figure 5.11: Strikes and maturities of all warrants under consideration.



risk aversion identifies several products in classes 2 and 3. We zoom in on some of these apparently rather attractive products in Section 5.4.3 below.

Table 5.12: Calls: Classifications for different model setups.

	BS		VG		NTS		AG		ED	
classification	V@R	γ_0	V@R	γ_0	V@R	γ_0	V@R	γ_0	V@R	γ_0
1	0	0	0	0	0	0	0	0	0	0
2	0	9	0	6	0	6	0	1	0	6
3	0	76	0	79	0	93	0	65	0	62
4	39	1,569	31	1,897	28	2,462	0	2,928	39	1,930
5	7,738	6,123	7,746	5,795	7,749	5,216	7,777	4,783	7,738	5,779

As a further robustness check, we illustrate how our results depend on the choice of the subjective discount factor α in Tables 5.13 and 5.14 for subjective annual discount rates of 10% and 5%, respectively. For each table the thresholds are determined with the same method as before, via calibrating to log-normally distributed payoffs with 90 days maturity.

Table 5.13: Calls: Classifications for subjective discount rate 10% p.a.

classification	limits	BS	VG	NTS	AG	ED
1	$59.9 < \gamma_0$	0	0	0	0	0
2	$24.0 < \gamma_0 \leq 59.9$	0	0	0	0	0
3	$13.0 < \gamma_0 \leq 24.0$	46	48	58	23	31
4	$8.2 < \gamma_0 \leq 13.0$	1,834	2,047	2,484	3,135	2,083
5	$\gamma_0 \leq 8.2$	5,897	5,682	5,235	4,619	5,663

Table 5.14: Calls: Classifications for subjective discount rate 5% p.a.

classification	limits	BS	VG	NTS	AG	ED
1	$68.8 < \gamma_0$	0	0	0	0	0
2	$50.4 < \gamma_0 \leq 68.8$	0	0	0	0	0
3	$27.2 < \gamma_0 \leq 50.4$	0	0	4	5	0
4	$17.2 < \gamma_0 \leq 27.2$	1,697	1,754	2,101	2,977	1,934
5	$\gamma_0 \leq 17.2$	6,080	6,023	5,672	4,795	5,843

We see that, as α increases (i.e., as the investor's impatience diminishes), the threshold values of γ_0 are shifted upward. This is due to the fact that a larger value of α implies a larger weight on the future payoff $X + \eta$ in (OEU), leading to a lower individual value of ρ^u for each product; this implies that the thresholds are shifted upward. The risk classifications, by contrast, are quite robust.

We next address put warrants. Here the overall picture, illustrated in Figure 5.13, looks rather different: The V@R and relative risk aversion ratings coincide, grouping all products into the highest risk class. With the ARMA-GARCH setup the only exception, relative risk aversion also indicates the highest possible risk throughout our sample. Accordingly, the risk classification in Table 5.15 is trivial.

Table 5.15: Puts: Classifications for different model setups.

	BS		VG		NTS		AG		ED	
classification	V@R	γ_0	V@R	γ_0	V@R	γ_0	V@R	γ_0	V@R	γ_0
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0
5	7,600	7,600	7,600	7,600	7,600	7,600	7,600	7,600	7,600	7,600

We wish to emphasize that this phenomenon should not be attributed to a shortcoming of either V@R or γ_0. The simple explanation is that our results are based on a two-year daily DAX® history with a positive drift rate. This implies that a put option, if it is priced fairly or subfairly, has a negative expected net payoff and should therefore be grouped into the highest risk class by any reasonable rating system. Concerning implied risk aversion, this can be seen formally as follows: By Corollary 4.12 we have for any net investment X

with $X_{\min} \leq 0$

$$\rho^u(X,\gamma) \geq -\alpha\mathbb{E}[X] \quad \text{for all } \gamma \geq 0.$$

Therefore, whenever $\mathbb{E}[X] \leq 0$, it follows that $\rho^u(X,\gamma) \geq 0$ for all $\gamma \geq 0$ and consequently $\gamma_0(X) = 0$. Nevertheless, put options that are "too cheap" may exhibit a positive expected return, and we demonstrate below that implied risk aversion is in fact able to identify such products.

5.4.2 Ranking call warrants with implied risk aversion and V@R

So far we have focused on the overall distribution of warrants into alternative risk classes. In the following we illustrate the qualitative advantages of implied risk aversion by analyzing in detail some typical scenarios with specific products. First, consider the subsample consisting of 5-day call options issued by Deutsche Bank with strikes ranging from 10,000 to 12,300 (the underlying traded around 11,830). Table 5.16 summarizes the implied risk aversion and V@R ratings of these products. V@R is increasing in the strike, and assigns maximum risk even to products that are in the money. From the retail client's perspective, this means that V@R does not offer any qualitative information for his investment decision beyond the products' obvious properties. Implied risk aversion, by contrast, is not monotone with respect to the strike: Calls deep in the money share roughly the same implied risk aversion. As the strike is increased, implied risk aversion falls slightly and then increases steeply and attains its maximum around the money, indicating the most attractive products. Implied risk aversion then falls again as the call moves out of the money.

To understand this behavior, note first that the net payoffs of the calls deep in the money are nearly identical and resemble a forward that knocks out only in a severe market crash. As the strike approaches the spot, the scenario of a total loss becomes more and more imminent, leading to a decrease in the product's rating; see, e.g., the call with strike 11,300 and Figure 5.14. If the strike is increased further, leverage makes the product more and more attractive, overcompensating the risk of a total loss and leading to a peak in implied risk aversion in the flat section of the bell curve in Figure 5.14.

The effect described above is amplified by the volatility skew, which for illustration is depicted in Figure 5.15. We wish to stress, however, that the same behavior persists in the absence of a volatility skew: In the last column of Table 5.16 we display the implied risk aversion of the call warrants under consideration under the assumption that they are priced in a Black-Scholes constant volatility setting.[6]

On a general note, short term at-the-money call warrants have the highest implied risk aversion which is, for example, reflected in the fact that strikes are in the range of 11,750 to 12,250 and time to maturity is at maximum 7 days among the 60 products with highest γ_0; see Table A.1.

5.4.3 Comparing products using implied risk aversion

Table 5.17 shows 4 far-out-of-the-money calls close to expiry, with the same maturities and strikes. Although their payoffs are identical, these products exhibit different prices. In

[6] Here, volatility is fixed at the implied volatility of XM2HAG.

Table 5.16: Implied risk aversion in the Black-Scholes model for 5-day call warrants issued by Deutsche Bank.

WKN	Strike	Ask	Ask-Vol	Underlying	V@R	γ_0	γ_0^{ATM-IV}
XM2H9D	10,000	18.32	0.5469	11,830.38	3,339	3.8707	3.9388
XM2H9M	10,400	14.32	0.4617	11,829.58	4,272	3.7237	3.9388
XM2H9V	10,800	10.32	0.3106	11,830.46	5,927	3.8966	3.9397
XM2H9Z	11,000	8.32	0.2884	11,829.11	7,352	3.6497	3.9493
XM2HA1	11,100	7.32	0.2762	11,827.80	8,356	3.4458	3.9826
XM2HA3	11,200	6.37	0.2589	11,831.40	9,681	3.3066	4.0635
XM2HA5	11,300	5.40	0.2509	11,830.66	10,000	2.9429	4.2255
XM2HA7	11,400	4.43	0.2242	11,831.50	10,000	3.2336	4.5354
XM2HA9	11,500	3.50	0.2158	11,829.89	10,000	3.1703	5.0840
XM2HAB	11,600	2.62	0.1974	11,831.96	10,000	4.2601	5.8402
XM2HAD	11,700	1.82	0.1875	11,830.93	10,000	5.5943	6.8032
XM2HAF	11,800	1.16	0.1820	11,829.02	10,000	7.1652	7.7712
XM2HAG	11,850	0.88	0.1783	11,827.33	10,000	8.4603	8.1599
XM2HAH	11,900	0.65	0.1713	11,831.82	10,000	11.0532	8.3605
XM2HAK	12,000	0.32	0.1667	11,831.29	10,000	12.2658	8.3810
XM2HAM	12,100	0.15	0.1686	11,830.42	10,000	9.6278	7.2299
XM2HAP	12,200	0.075	0.1769	11,829.15	10,000	5.6930	5.3404
XM2HAR	12,300	0.035	0.1829	11,829.13	10,000	3.2332	3.5792

this context, we wish to point out that not only is implied risk aversion γ_0 able to identify the most attractive (i.e., cheapest) product (AP6RRY), but γ_0 also provides an indication which type of investor may be interested in such a warrant (within, say, the Black-Scholes model): Featuring an implied risk aversion of 0.1451, it is only for the least risk-averse investors. Furthermore, if AP6RRY were not in the sample, implied risk aversion would clearly indicate that none of the remaining products, which all exhibit zero implied risk aversion, is attractive for a risk-averse retail investor.

Table 5.17: Implied risk aversion of 7-day call warrants with strike 12,600.

WKN	Issuer	Ask	Underlying	γ_0^{BS}	γ_0^{VG}	γ_0^{NTS}	γ_0^{AG}	γ_0^{ED}
CC8Z2S	Citigroup	0.031	11,830.85	0	0	0	1.6735	0
CN0R7V	Commerzbank	0.025	11,829.26	0	0.0551	0.0139	1.9035	0
GL15K9	Goldman Sachs	0.031	11,830.40	0	0	0	1.6670	0
AP6RRY	Int. Brokers	0.018	11,830.32	0.1451	0.3780	0.2778	2.2155	0.1100

In a similar vein, Table 5.18 displays implied risk aversion parameters for 5-day at-the-money call warrants. Implied risk aversion clearly identifies products issued by UBS as the most attractive ones. Comparing with prices of identical warrants issued by Deutsche Bank (DB) or Goldman Sachs (GS), it becomes clear that the former are indeed up to 10% cheaper than the latter.

Table 5.18: Implied risk aversion of 5-day at-the-money call warrants.

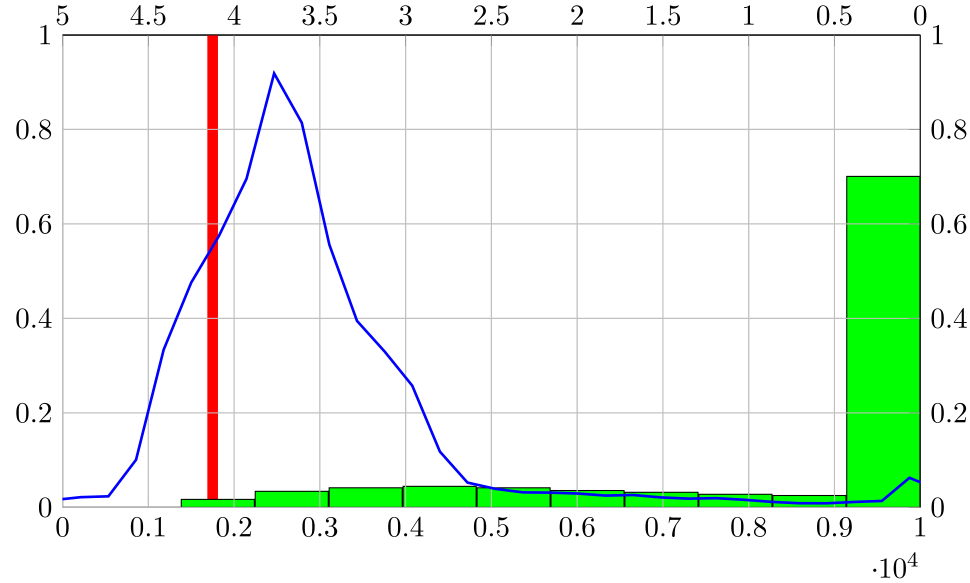
WKN	Issuer	Strike	Ask	Underlying	γ_0^{BS}	γ_0^{VG}	γ_0^{NTS}	γ_0^{AG}	γ_0^{ED}
XM2HAG	DB	11,850	0.88	11,827.33	8.4603	7.9492	9.2904	7.0981	7.3603
UZ698B	UBS	11,850	0.81	11,829.76	13.2928	12.7038	14.4364	11.2965	12.0095
XM2HAH	DB	11,900	0.65	11,831.82	11.0532	10.2653	11.4942	8.6166	9.5661
GL15JU	GS	11,900	0.67	11,829.52	9.1773	8.4441	9.5349	7.0891	7.7896
UZ7M77	UBS	11,900	0.59	11,820.62	15.0835	14.2694	15.6127	11.9888	13.5410
XM2HAJ	DB	11,950	0.47	11,830.66	11.5091	10.5665	11.3985	8.5966	9.8672
UZ68WB	UBS	11,950	0.41	11,828.86	15.4231	14.5386	15.3694	11.8845	13.9345
XM2HAK	DB	12,000	0.32	11,831.29	12.2658	11.3418	11.8613	9.2192	10.8105
GL15JW	GS	12,000	0.327	11,831.26	11.7561	10.8253	11.3453	8.8088	10.2802
UZ62LW	UBS	12,000	0.28	11,829.15	14.6950	13.8157	14.3215	11.2509	13.3469
XM2HAL	DB	12,050	0.22	11,829.58	10.9613	10.0404	10.3126	8.3815	9.5545
UZ7LV7	UBS	12,050	0.18	11,828.89	13.6851	12.8125	13.0584	10.5945	12.3940
XM2HAM	DB	12,100	0.15	11,830.42	9.6278	8.7600	8.8946	7.6674	8.3252
GL15JY	GS	12,100	0.145	11,831.93	10.1782	9.3134	9.4372	8.0739	8.8888
UZ6U6M	UBS	12,100	0.11	11,829.38	12.3819	11.5603	11.6264	9.8478	11.1888

This picture is confirmed if we consider put options, see Table 5.19. While, as discussed above, none of the put warrants looks particularly attractive, some of those issued by UBS may be attractive to some investors, and are clearly (within this sample, in this dataset) the most attractive products.

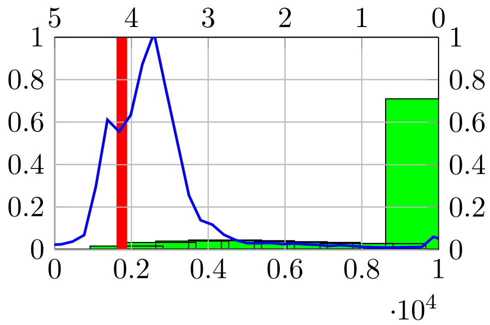
Table 5.19: Implied risk aversion of 5-day at-the-money put warrants.

WKN	Issuer	Strike	Ask	Underlying	γ_0^{BS}	γ_0^{VG}	γ_0^{NTS}	γ_0^{AG}	γ_0^{ED}
XM2HUV	DB	11,400	0.11	11,829.52	0	0	0	0.3770	0
GL15LV	GS	11,400	0.134	11,829.08	0	0	0	0.0508	0
UZ66RH	UBS	11,400	0.074	11,818.24	0	0.0002	0.0368	1.7827	0
XM2HUZ	DB	11,600	0.31	11,831.63	0	0	0	0.0290	0
GL15LZ	GS	11,600	0.341	11,831.02	0	0	0	0	0
UZ68VB	UBS	11,600	0.26	11,823.39	0	0.0381	0.0666	0.7609	0
XM2HV3	DB	11,800	0.86	11,827.6	0	0	0	0	0
GL15M3	GS	11,800	0.91	11,828.79	0	0	0	0	0
UZ697A	UBS	11,800	0.82	11,820.37	0.0297	0.0296	0.0023	0.0310	0.0049

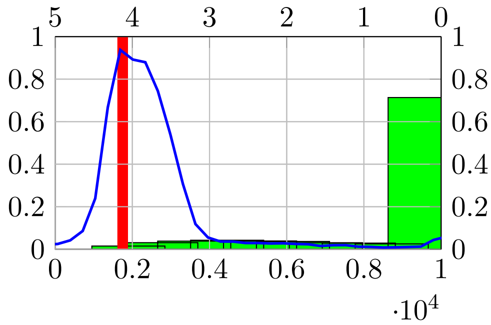
Figure 5.12: Calls: Distribution of V@R vs. distribution of γ_0 for the considered models.



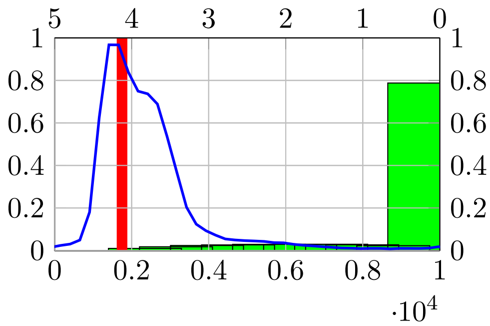
(a) Black-Scholes



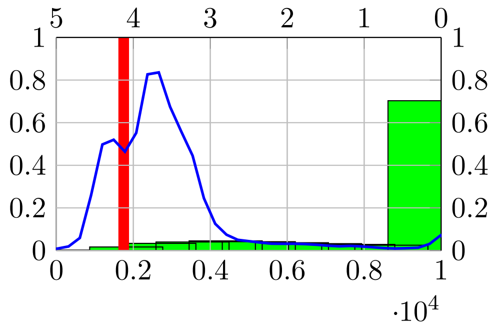
(b) Variance Gamma



(c) NTS

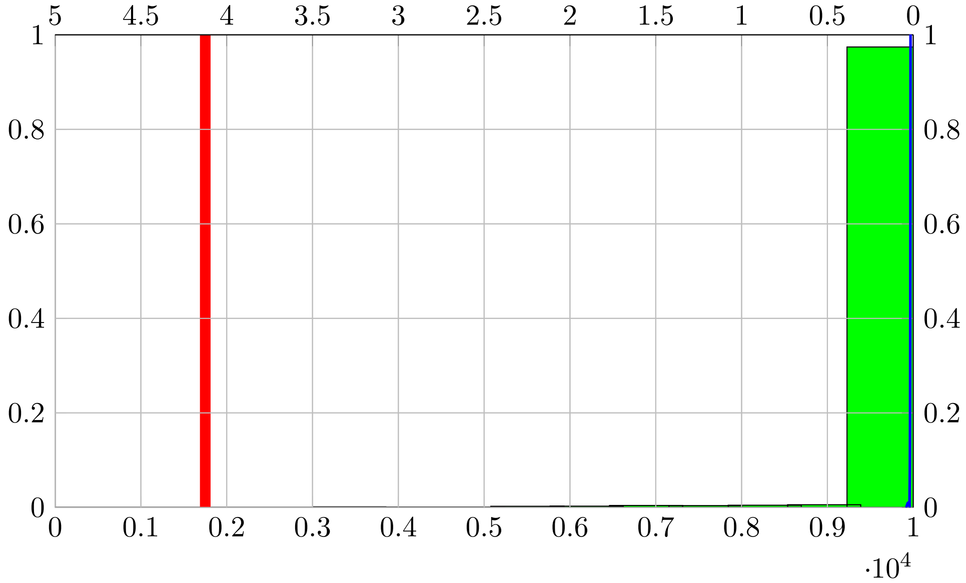


(d) ARMA-GARCH

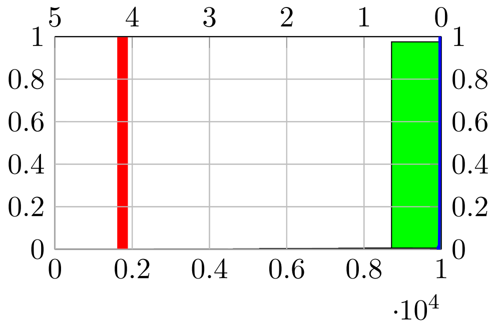


(e) Empirical distribution

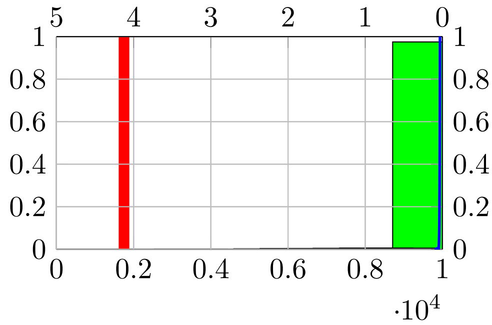
Figure 5.13: Puts: Distribution of V@R vs. distribution of γ_0 for the considered models.



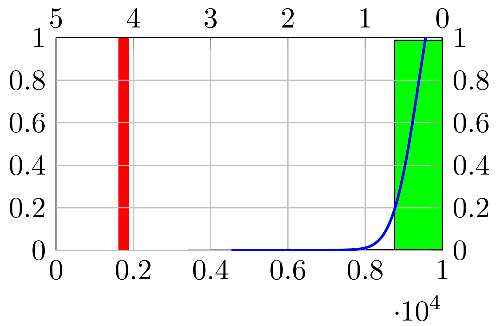
(a) Black-Scholes



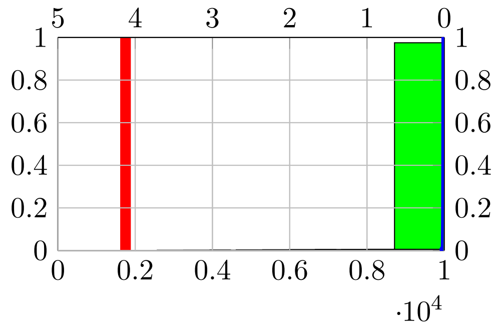
(b) Variance Gamma



(c) NTS



(d) ARMA-GARCH



(e) Empirical distribution

Figure 5.14: Estimated distribution of the 5-day-DAX®. As one can see, only from strikes around 11,000 the density function starts to rise significantly.

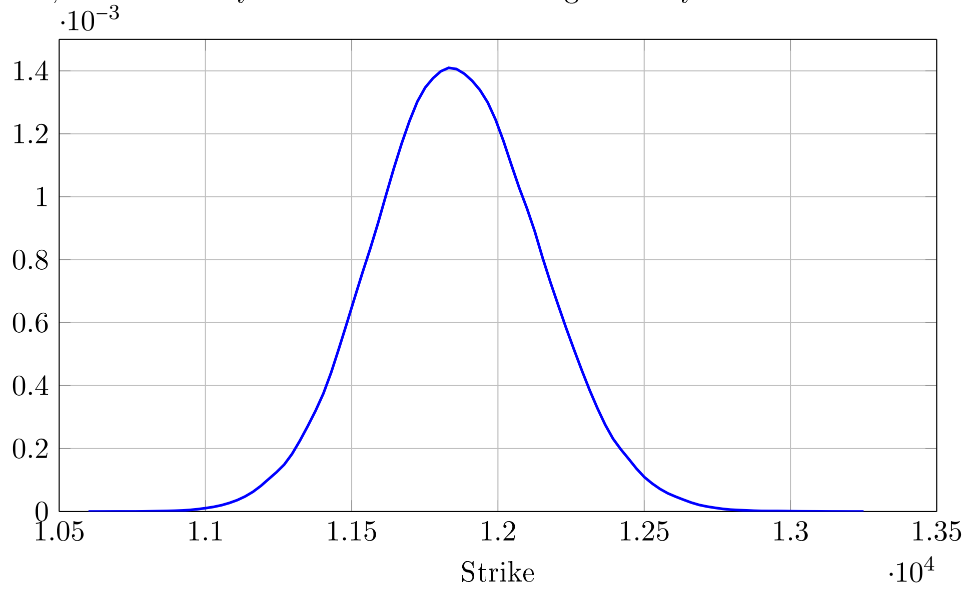
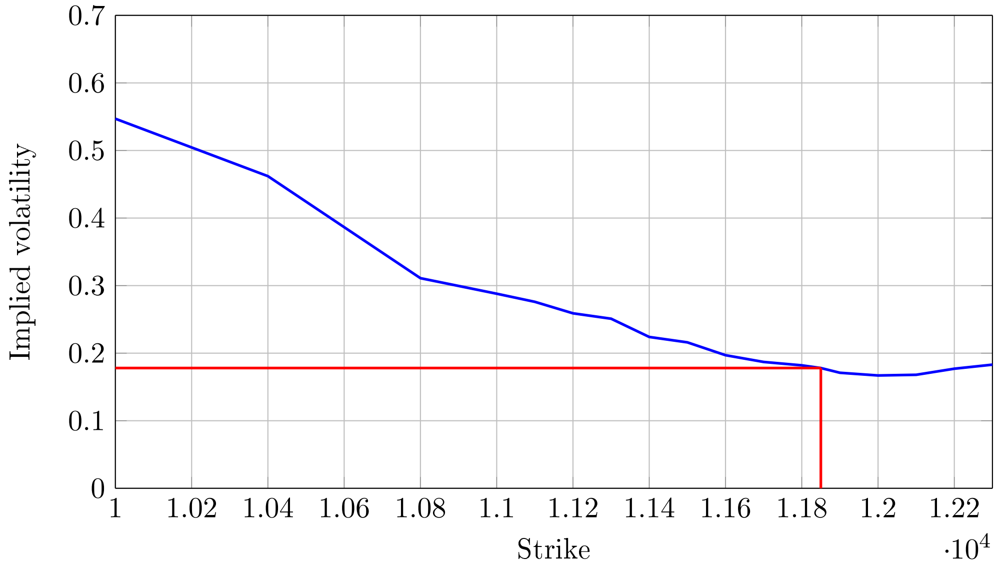


Figure 5.15: Implied volatilities for the Deutsche Bank call warrants of Table 5.16. Red line marks approximate at-the-money volatility.



CHAPTER 6

Risk measures on the space of distributions and time consistency of risk measures

The previous chapters deal with the standard case of risk measures on the space of random variables. If, however, the outcome of risk measures only depends on the distribution of the considered random variables, these risk measures can also be interpreted as functionals on the space of distributions. In this chapter we consider risk measures on distributions and point out some fundamental attributes of these functionals with a particular focus on optimal expected utility risk measures. Risk measures on the space of probability distributions have been studied by Acciaio and Svindland [2], Drapeau and Kupper [54], Föllmer and Weber [72], Fritelli, Maggis and Peri [75] and Weber [136], [137], among others.

The interrelation of conditional risk measures at different time steps is a crucial property of dynamic risk measures; see Section 2.3 and Section 2.4. In this chapter we point out which properties static risk measures should fulfill so that it is reasonable to build up conditional risk measures on them. We introduce optimal expected utility risk measures for distributions and show that these are concave on the space of distributions - a property which is further discussed in [2].

In Section 6.1 we set up the definition of risk measures on the space of distributions, study how OEU and other risk measures react on extreme event scenarios and give some exemplary results on the convexity of acceptance and rejection sets. In Section 6.2 we recall some of the results from [137] on connections between risk measures on distributions in static time and properties of related dynamic risk measures. Based on this, in Section 6.3 we consider time consistency properties of dynamic risk measures with time-dependent parameters and thereby slightly generalize the results given in [137].

6.1 Concave risk measures on the space of distributions

Following [137], from now on we fix a probability measure P on $(\Omega,\mathcal{F})$ and assume that $(\Omega,\mathcal{F},P)$ is an atomless standard Borel probability space such that any probability distribution μ on $\mathbb{R}$ can be represented as the distribution of a random variable X on $(\Omega,\mathcal{F},P)$. For this chapter we assume $I = \{t_0 = 0, t_1, ..., t_N = T\}$, i.e., we only consider the case of DDRMs.

6.1 Definition. The *distribution* μ of a random variable $X \in \mathcal{X}$ is the image of P under X. For a given distribution $\mu \in \mathcal{M}_{1,c}$ we write

$$X \sim \mu \text{ if } \mu = P \circ X^{-1}.$$

By $\mathcal{M}_{1,c}$ we denote the space of probability measures on $\mathbb{R}$ with compact support:

$$\mathcal{M}_{1,c} = \{\mu : X \sim \mu,\ X \in \mathcal{X}\}.$$

The elements of $\mathcal{M}_{1,c}$ are also called *lotteries*.

6.2 Definition. A risk measure ρ is called *law-invariant* if $\rho(X_1) = \rho(X_2)$ whenever X_1 and X_2 have the same distribution under P.

Law-invariant risk measures of random variables X on some probability space $(\Omega,\mathcal{F},P)$ only depend on the distribution μ of X under P and can therefore also be understood as functionals on the space of distributions. Clearly, value at risk, average value at risk and utility-based shortfall risk are law-invariant; see [72, page 18]. We show that also OEU is law-invariant.

6.3 Lemma. *ρ^u is a law-invariant risk measure.*

Proof. Let $X,Y \in \mathcal{X}$ such that X and Y have the same distribution, i.e.,

$$P(X > t) = P(Y > t) \quad \text{for all } t \in \mathbb{R}.$$

Then, we have that for all $u \in \mathcal{U}$

$$\mathbb{E}[u(X)] = \int u(\xi)P(X \in d\xi) = \int u(\xi)P(Y \in d\xi) = \mathbb{E}[u(Y)].$$

It clearly follows that if X and Y have the same distribution, then also $X + \eta$ and $Y + \eta$ have the same distribution for any $\eta \in \mathbb{R}$, thus we get

$$-\beta\eta + \alpha u^{-1}(\mathbb{E}[u(X+\eta)]) = -\beta\eta + \alpha u^{-1}(\mathbb{E}[u(Y+\eta)]) \quad \text{for all } \eta \in \mathbb{R},\ 0 < \alpha < \beta \in \mathbb{R},$$

and hence

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

□

The interpretation of law-invariant risk measures becomes particularly helpful in the context of *mixture distributions* which means the following: Suppose $X_1, X_2 \in \mathcal{X}$ are financial positions with known distributions μ_1, μ_2. If we add random weights $0 \leq \lambda \leq 1$ to choose μ_1 and $1 - \lambda$ to choose μ_2 (think of an independent Bernoulli random variable

Y which takes the value 0 with probability λ and the value 1 with probability $1 - \lambda$) and consider the compound position X with

$$X(\omega) = \begin{cases} X_1(\omega), \text{ if } Y(\omega) = 0, \\ X_2(\omega), \text{ if } Y(\omega) = 1 \end{cases} \qquad \text{for all } \omega \in \Omega,$$

then μ is the *mixture distribution* of μ_1 and μ_2 with

$$X \sim \mu, \quad \mu = \lambda\mu_1 + (1 - \lambda)\mu_2.$$

The question arises how the risks of X_1 and X_2 are related to the risk of X. For this purpose, we introduce a partial order on $\mathcal{M}_{1,c}$ by first-order stochastic dominance, i.e., $\mu_1 \leq \mu_2$ if

$$\int f d\mu_1 \leq \int f d\mu_2 \quad \text{for all increasing functions } f : \mathbb{R} \to \mathbb{R}.$$

Remark that there is a general difference between μ on the one hand and the law of the state-wise convex combination $\lambda X_1 + (1 - \lambda)X_2$ on the other hand which is the reason why we need to be careful, when translating features of risk measures on the space of random variables to risk measures on the space of distributions. Risk measures on the space of distributions are defined as follows:

6.4 Definition. A function $\psi : \mathcal{M}_{1,c} \to \mathbb{R}$ is called a *risk measure on the space of distributions* if it satisfies the following conditions for all $\mu, \mu_1, \mu_2 \in \mathcal{M}_{1,c}$:

(M) *Monotonicity*: $\psi(\mu_1) \geq \psi(\mu_2)$ if $\mu_1 \leq \mu_2$.

(CI) *Cash-invariance*: $\psi(\tilde{T}_m\mu) = \psi(\mu) - \beta m$ for $m \in \mathbb{R}$, $0 < \beta$,

where the translation operator $\tilde{T}_m$ is given by

$$(\tilde{T}_m\mu)(\cdot) = \mu(\cdot - m).$$

We introduce optimal expected utility risk measures on $\mathcal{M}_{1,c}$:

6.5 Definition. For any $u \in \mathcal{U}$ and any distribution $\mu \in \mathcal{M}_{1,c}$, the *optimal expected utility risk measure* of μ is defined by the map $\psi^u(\mu) : \mathcal{M}_{1,c} \to \mathbb{R}$,

$$\psi^u(\mu) := \rho^u(X) = -\sup_{\eta \in \mathbb{R}} \left\{ -\beta\eta + \alpha u^{-1} \left(\mathbb{E}^\mu \left[u\left(X + \eta \right) \right] \right) \right\},$$

where $X \in \mathcal{X}$ is such that $X \sim \mu$.

From [137, page 421] we know that, since ρ^u is a law-invariant risk measure on $\mathcal{X}$, $\psi^u(\mu) := \rho^u(X)$ for some $X \sim \mu$ defines a risk measure on $\mathcal{M}_{1,c}$ in the sense of Definition 6.4. However, for the sake of completeness, we explicitely show that ψ^u fulfills (M) and (CI) for any $\mu \in \mathcal{M}_{1,c}$.

6.6 Theorem. *$\psi^u(\mu)$ is a risk measure on the space of distributions.*

Proof. (M) If $X_1 \sim \mu_1$ and $X_2 \sim \mu_2$ with $\mu_1 \leq \mu_2$, we know that for $u \in \mathcal{U}$:

$$\mathbb{E}[u(X_1 + \eta)] \leq \mathbb{E}[u(X_2 + \eta)],$$

for all $\eta \in \mathbb{R}$, and since u^{-1} is an increasing function consequently

$$\psi^u(\mu_1) \geq \psi^u(\mu_2).$$

(CI) For $m \in \mathbb{R}$, if $X \sim \tilde{T}_m\mu$, then $X - m \sim \mu$. Thus,

$$\begin{aligned}
\psi^u(\tilde{T}_m\mu) &= -\sup_{\eta\in\mathbb{R}} \left\{-\beta\eta + \alpha u^{-1}\left(\mathbb{E}\left[u\left(X+\eta\right)\right]\right)\right\} \\
&= -\sup_{(\eta-m)\in\mathbb{R}} \left\{-\beta(\eta-m) + \alpha u^{-1}\left(\mathbb{E}\left[u\left(X+\eta-m\right)\right]\right)\right\} \\
&= -\sup_{\eta\in\mathbb{R}} \left\{-\beta\eta + \alpha u^{-1}\left(\mathbb{E}\left[u\left((X-m)+\eta\right)\right]\right)\right\} - \beta m \\
&= \psi^u(\mu) - \beta m.
\end{aligned}$$

□

Following [78, Section 5], the objective of Example 6.7 is to compare the sensitivity of value at risk, average value at risk and OEU with respect to extreme events.

6.7 Example. Consider an extreme event scenario where the *normal case* is a floored lognormal distribution

$$X_1 = -(Z_1 \wedge 11), \quad \text{where} \quad Z_1 \sim \mu_1 = \ln\mathcal{N}(0,0.1),$$

and the *extreme case* is modeled by a floored normal distribution

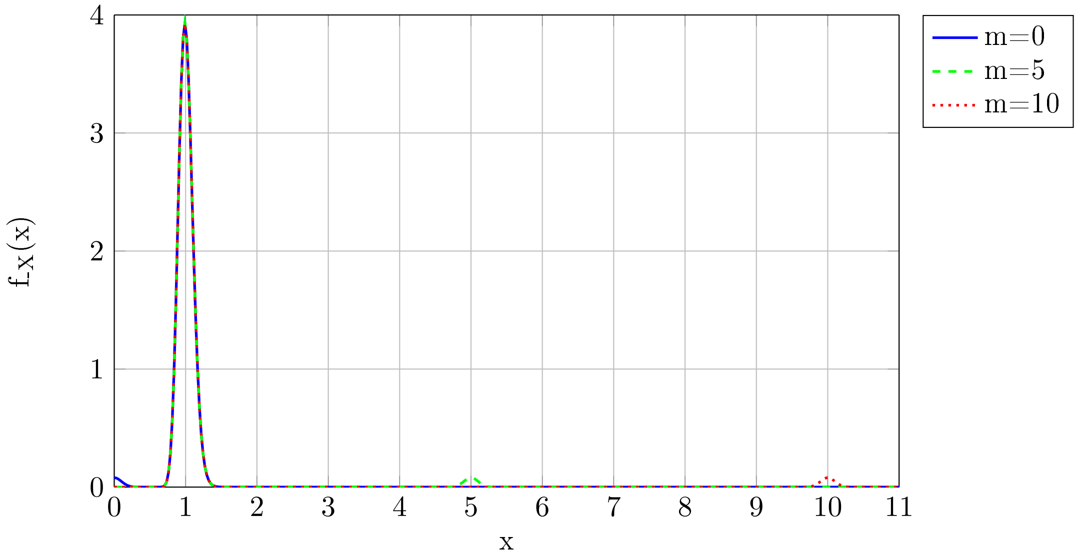
$$X_2 = -(Z_2 \wedge 11), \quad \text{where} \quad Z_2 \sim \mu_2 = \mathcal{N}(m,0.1), \quad m \in [0,10]$$

and we consider the mixture distribution

$$X(\omega) = \begin{cases} X_1(\omega), \text{ if } Y(\omega) = 0, \\ X_2(\omega), \text{ if } Y(\omega) = 1 \end{cases} \qquad \text{for all } \omega \in \Omega,$$

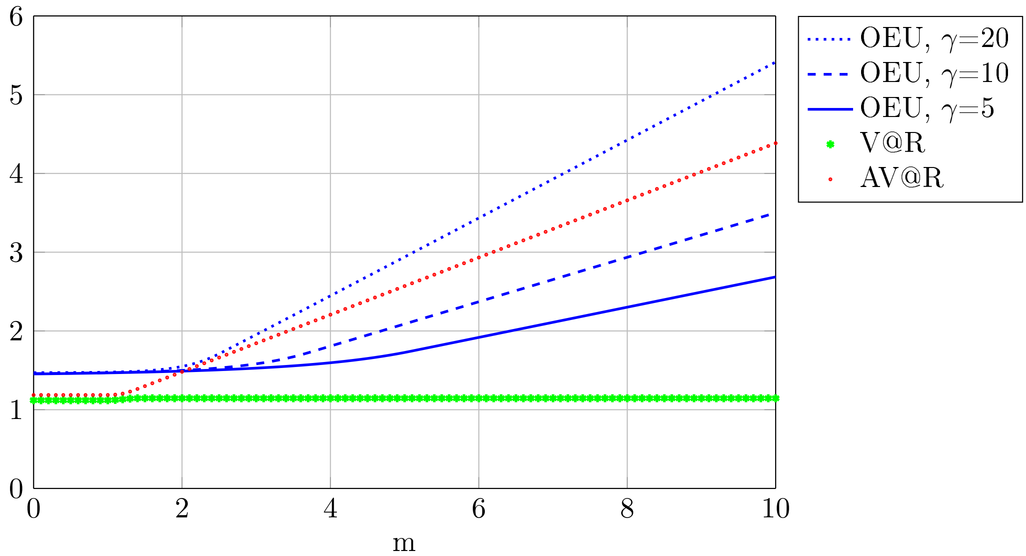
where Y is an independent Bernoulli random variable which takes the value 0 with probability 0.98 and the value 1 with probability 0.02. The distribution of $-X$ looks as follows:

Figure 6.1: Distribution of -X.



We want to compare how sensitively $V@R$, $AV@R$ and OEU react to changes in the mean of the extreme event scenario m. We choose the level $\lambda = 0.05$, ρ^u is evaluated with $u(x) = \frac{1}{1-\gamma}\left(x^{1-\gamma} - 1\right)$ for $\gamma = 5{,}10{,}20$ and we choose $\alpha = 0.9$, $\beta = 0.95$.

Figure 6.2: Sensitivity to extreme risks.



We notice that $V@R$ does only reflect the riskiness of extreme losses for small choices of m since the probability of extreme losses is smaller than the chosen level λ. OEU and $AV@R$, however, do not have this drawback but indicate potential extreme risks. This example clearly demonstrates that OEU recognizes extreme downside risks although it is not a pure downside risk measure as AV@R; see also Section 5.2.1. In particular, we note that OEU attains greater values than AV@R for m being smaller than 2 which is due to the fact that OEU always refers to the possible worst case scenario which in this setting is set to $X_{\min} = -11$; see Remark 4.23.

Next we consider convexity and concavity of risk measures on distributions. As shown in [136, Lemma 2.3.8], the notion of convexity of risk measures on $\mathcal{M}_{1,c}$ is well defined. Whereas convexity of risk measures on random variables is a widely demanded and recognized property, see Definition 2.1 and the subsequent remarks on the diversification effect, this is no more obvious on the level of distributions; see [54, Remark 2]. An alternative notion of diversification on the space of distributions must rather be reconsidered. A convex combination of distributions $\mu = \lambda\mu_1 + (1-\lambda)\mu_2$ corresponds to an additional randomization which can be interpreted as an additional factor of risk coming from an independent toss with probabilities λ and $(1-\lambda)$ deciding which lottery is chosen: Acciaio and Svindland [2] show that given a law-invariant convex risk measure on the space of random variables, for the respective risk measure on the space of distributions "if, at all, concavity holds true", [2, page 54]. Weber and Schmidt [138] argue that, from a descriptive perspective, a mixture distribution μ might be rejected by investors even if both, μ_1 and μ_2, are considered acceptable which relates to concavity on the space of distributions. The following theorem shows that OEU is concave on $\mathcal{M}_{1,c}$.

6.8 Theorem. *$\psi^u(\mu)$ is concave on $\mathcal{M}_{1,c}$.*

Proof. Let $\mu_1,\mu_2 \in \mathcal{M}_{1,c}$ and, for some $0 \le \lambda \le 1$, consider $X, X_1, X_2 \in \mathcal{X}$ such that $X \sim \lambda\mu_1 + (1-\lambda)\mu_2$, $X_1 \sim \mu_1$, $X_2 \sim \mu_2$. Note that, for any $\eta \in \mathbb{R}$

$$\begin{aligned}\mathbb{E}[u(X+\eta)] &= \int u(\xi)P(X+\eta \in d\xi)\\ &= \lambda\int u(\xi)P(X_1+\eta\in d\xi) + (1-\lambda)\int u(\xi)P(X_2+\eta\in d\xi)\\ &= \lambda\mathbb{E}[u(X_1+\eta)] + (1-\lambda)\mathbb{E}[u(X_2+\eta)].\end{aligned}$$

Thus,

$$\begin{aligned}\psi^u(\lambda\mu_1+(1-\lambda)\mu_2) &= -\sup_{\eta\in\mathbb{R}}\left\{-\beta\eta + \alpha u^{-1}\left(\mathbb{E}\left[u\left(X+\eta\right)\right]\right)\right\}\\ &= -\sup_{\eta\in\mathbb{R}}\left\{-\beta\eta + \alpha u^{-1}\left(\lambda\mathbb{E}\left[u\left(X_1+\eta\right)\right] + (1-\lambda)\mathbb{E}\left[u\left(X_2+\eta\right)\right]\right)\right\}\\ &\ge -\sup_{\eta\in\mathbb{R}}\left\{-\beta\eta + \lambda\alpha u^{-1}\left(\mathbb{E}\left[u\left(X_1+\eta\right)\right]\right) + (1-\lambda)\alpha u^{-1}\left(\mathbb{E}\left[u\left(X_2+\eta\right)\right]\right)\right\}\\ &= -\sup_{\eta\in\mathbb{R}}\Big\{\lambda\left(-\beta\eta + \alpha u^{-1}\left(\mathbb{E}\left[u\left(X_1+\eta\right)\right]\right)\right)\\ &\qquad\qquad + (1-\lambda)\left(-\beta\eta + \alpha u^{-1}\left(\mathbb{E}\left[u\left(X_2+\eta\right)\right]\right)\right)\Big\}\\ &\ge \lambda\left(-\sup_{\eta\in\mathbb{R}}\left\{-\beta\eta + \alpha u^{-1}\left(\mathbb{E}\left[u\left(X_1+\eta\right)\right]\right)\right\}\right)\\ &\qquad\qquad + (1-\lambda)\left(-\sup_{\eta\in\mathbb{R}}\left\{-\beta\eta + \alpha u^{-1}\left(\mathbb{E}\left[u\left(X_2+\eta\right)\right]\right)\right\}\right)\\ &= \lambda\psi^u(\mu_1) + (1-\lambda)\psi^u(\mu_2),\end{aligned}$$

where we have used in the third step that u^{-1} is convex on the domain of u; it can be easily verified that the second derivative of u^{-1} is positive at all points of the domain of u for any $u \in \mathcal{U}$. □

Due to [2, Example 7 (2.)], AV@R is a concave risk measure on distributions. Value at risk is quasi-convex on distributions, as shown in [54, Example 10], but since it is not quasi-convex on the space of random variables, Acciaio and Svindland follow from [2, Lemma 3] that "there are payoff profiles X,Y such that every expected utility agent prefers X to Y, but under V@R the profile Y is strictly less risky than X"; [2, page 58], we also refer to Section 5.2.2.

Similarly to Definition 2.3, we introduce acceptance and rejection sets of risk measures on distributions:

6.9 Definition. Let $\psi : \mathcal{M}_{1,c} \to \mathbb{R}$ be any risk measure on distributions. We call the set of all distributions with negative risk

$$\mathcal{N}_\psi = \{\mu \in \mathcal{M}_{1,c} : \psi(\mu) \leq 0\}$$

the *acceptance set* of ψ and interpret any $\mu \in \mathcal{N}_\psi$ as *acceptable*. Analogously, we call the set of all distributions with strictly positive risk

$$\mathcal{N}_\psi^c = \{\mu \in \mathcal{M}_{1,c} : \psi(\mu) > 0\}$$

the *rejection set* of ψ and interpret any $\mu \in \mathcal{N}_\psi^c$ as *rejectable*.

As shown in [72], due to cash-invariance, any risk measure ψ is quasi-convex on $\mathcal{M}_{1,c}$ if and only if the acceptance set $\mathcal{N}_\psi$ is convex. Similarly, ψ is quasi-concave on $\mathcal{M}_{1,c}$ if and only if the rejection set $\mathcal{N}_\psi^c$ is convex; see also [15, Lemma 2.2]. Thus it follows from Theorem 6.8 that for ψ^u, whenever two distributions $\mu_1, \mu_2 \in \mathcal{M}_{1,c}$ are rejectable and $\lambda \in [0,1]$ is some probability, then the compound probability measure $\lambda\mu_1 + (1-\lambda)\mu_2$ is also rejectable for OEU. We continue with an example of two financial positions $X_1, X_2 \in \mathcal{X}$ which are acceptable for AV@R, but whose compound position may not be accepted when evaluated with AV@R.

6.10 Example. Let $X_1, X_2 \in \mathcal{X}$ with the following payout profiles:

$$X_1 = \begin{cases} -10, & \text{with } p = 0.01, \\ 3, & \text{with } p = 0.99 \end{cases}, \quad X_2 = \begin{cases} -1, & \text{with } p = 0.02, \\ 1, & \text{with } p = 0.98 \end{cases}.$$

The average value at risk for level $\lambda = 0.05$ of X_1 and X_2 is

$$AV@R^{0.05}(X_1) = -\frac{1}{0.05}(-10 \cdot 0.01 + 3 \cdot 0.04) = -0.4,$$
$$AV@R^{0.05}(X_2) = -\frac{1}{0.05}(-1 \cdot 0.02 + 1 \cdot 0.03) = -0.2.$$

We now assume Y to be an independent random variable which takes the values 0 and 1 with probability $\frac{1}{2}$ and the compound position X of X_1 and X_2 to be defined as

$$X(\omega) = \begin{cases} X_1(\omega), \text{ if } Y(\omega) = 0, \\ X_2(\omega), \text{ if } Y(\omega) = 1 \end{cases} \qquad \text{for all } \omega \in \Omega.$$

The payoff of X is:

$$X = \begin{cases} -10, & \text{with } p = 0.005, \\ -1, & \text{with } p = 0.01, \\ 1, & \text{with } p = 0.49, \\ 3, & \text{with } p = 0.495 \end{cases}$$

and the average value at risk is

$$AV@R^{0.05}(X) = -\frac{1}{0.05}\left(-10 \cdot 0.005 - 1 \cdot 0.01 + 1 \cdot 0.035\right) = 0.5.$$

For an economic interpretation of the previous example think of X_1 as an investment in a speculative financial product with high possible losses and gains and X_2 as an investment in a more conservative product with relatively small potential profits and losses. The compound position of X_1 and X_2 shifts the payout in a way that the (still achievable) high profit of 3 is not affecting the risk calculation of the investment anymore which is due to the fact that AV@R only considers the downside risk of a position below the level λ. We conclude that AV@R is concave but not quasi-convex on $\mathcal{M}_{1,c}$.

Even though we cited Weber and Schmidt [138] in the context of concavity of risk measures on distributions, they also point out that, from a normative perspective, they find it more reasonable to assume that if any two distributions μ_1 and μ_2 are both acceptable (rejectable) then also the convex combination of these distributions should be acceptable (rejectable) and call this property *invariance under randomization*. Weber [136], [137] showed that under some mild regularity conditions, a law-invariant risk measure ψ on the space of distributions must be a utility-based shortfall risk measure for some increasing loss function l if both, the acceptance set and the rejection set of ψ are convex.

6.2 (Weak) time consistency of dynamic risk measures with time-independent risk aversion parameters

As explained in Section 2.4, a crucial property of dynamic risk measures is the relation of consecutive conditional risk measures to each other. In this section we recall some of the notable results from [137] which relate the convexity of acceptance and rejection sets of static risk measures on distributions to consistency properties of the respective dynamic risk measures. Let us first give a definition of law-invariance of dynamic risk measures.

6.11 Definition (Definition 4.3 from [137]). A dynamic risk measure $(\rho_t)_{t\in I}$ is called *law-invariant*, if there exists a measurable mapping $H_t : \mathcal{M}_{1,c} \to \{0,1\}$ such that for all $X \in \mathcal{X}$,

$$\mathbf{1}_{\{\rho_t(X)\leq 0\}} = H_t(\mathcal{L}(X \mid \mathcal{F}_t)).$$

Next we establish the definition of local measure convexity which becomes relevant in the upcoming theorems.

6.12 Definition (Definition 4.5 from [137]). Let $\mathcal{C}$ be a measurable subset of $\mathcal{M}_{1,c}$. We say that $\mathcal{C}$ is *locally measure convex* if for all $a \in \mathbb{R}$ and any probability measure ι on $\mathcal{C} \cap \mathcal{M}_1([-a,a])$ the mixture $\int_{\mathcal{C}} \nu\iota(d\nu)$ is again an element of $\mathcal{C}$, where $\mathcal{M}_1([-a,a])$ denotes the set of probability measures supported on the interval $[-a,a]$.

Note that $\int \nu\iota(d\nu)$ is a generalization of the mixture of two distributions as introduced at the beginning of Section 6.1.

6.13 Theorem (Theorem 4.2 from [137]). *Let ψ be a static risk measure, $\mathcal{N} \subseteq \mathcal{M}_{1,c}$ be its acceptance set, and $(\rho_t)_{t\in I}$ be the law-invariant dynamic risk measure defined by*

$$\rho_t(X) = \beta_t^T \cdot \psi(\mathcal{L}(X \mid \mathcal{F}_t)) \quad P\text{-}a.s. \tag{DynRM}$$

If $\mathcal{N}$ is locally measure convex, then $(\rho_t)_{t\in I}$ is acceptance consistent. If $\mathcal{N}^c$ is locally measure convex, then $(\rho_t)_{t\in I}$ is rejection consistent.

Remember that β_t^T denotes the value of a default-free bond at time t with face value 1 and maturity time T. Acceptance (rejection) consistency of dynamic risk measures means that acceptable (rejectable) financial positions at any time $t+1$ should also be accepted (rejected) at the earlier time t; see Definition 2.14.

If the underlying probability space is rich enough, the result from Theorem 6.13 can be strengthened.

6.14 Definition (Definition 4.6 from [137]). The filtered probability space is called *sequentially rich* if there exist both a unif(0,1)-distributed random variable independent of $\mathcal{F}_{T-1}$, and a unif(0,1)-distributed, $\mathcal{F}_{T-1}$-measurable random variable independent of $\mathcal{F}_{T-2}$.

6.15 Theorem (Theorem 4.3 from [137]). *Let the underlying probability space be sequentially rich, and assume that the dynamic risk measure $(\rho_t)_{t\in I}$ is represented as in* (DynRM). *Then $(\rho_t)_{t\in I}$ is acceptance consistent if and only if $\mathcal{N}$ is locally measure convex. Analogously, $(\rho_t)_{t\in I}$ is rejection consistent if and only if $\mathcal{N}^c$ is locally measure convex.*

Weber showed that if $(\rho_t)_{t\in I}$ is weakly time consistent, i.e., it is both acceptance and rejection consistent, then it can be represented by (DynRM) for a unique static risk measure ψ. In particular, if $(\rho_t)_{t\in I}$ is weakly time consistent, its characteristics over time can only change slightly by the multiplicative discount factor β_t^T but no shift or change of risk parametrization can happen.

6.16 Remark. Kupper and Schachermayer [98] show that the only dynamic risk measure which is law-invariant (in the sense that $\rho_0(X) = \rho_0(Y)$ whenever X and Y have the same distribution), time consistent and *relevant* (that is $\rho_0(-\varepsilon\mathbf{1}_A) > 0$ for all $A \in \mathcal{F}$ and all $\varepsilon > 0$) is the entropic risk measure with time-independent parameters. Clearly, any risk measure which is fully supported in the sense of Definition 5.1 is also relevant. Thus dynamic optimal expected utility risk measures can be considered as a natural expansion of the entropic risk measure in this respect: OEU are law-invariant and relevant, and even though they are not generally time consistent, in the special case of $u(x) = \frac{1}{\gamma}(1-\exp(-\gamma x))$ we obtain (a variant of) the time consistent entropic risk measure; see Example 4.10 (e).

6.3 (Weak) time consistency of dynamic risk measures with time-dependent risk aversion parameters

In this section we particularly study consistency properties of dynamic risk measures with time-dependent parameters. This is an important generalization of Section 6.2 since, as shown by Pollak [113] and Strotz [128], among others, preferences of an economic agent can change over time which then implies time-dependent risk aversion parameters of dynamic risk measures. We start this section with the example of dynamic average value at risk with time-dependent levels $(\lambda_t)_{t\in I}$.

6.17 Example. As shown in [1, Example 1.38 (2.)], $(AV@R_t^{\lambda_t})_{t\in I}$ is acceptance consistent if

$$\lambda_{t+1} \le \lambda_t \operatorname*{ess\,inf}_{Q\in\mathcal{R}_t} \mathbb{E}\left[\frac{dQ}{dP}|\mathcal{F}_{t+1}\right] \quad \text{for all } t \in I,$$

where $\mathcal{R}_t$ is some convex subset of $\mathcal{Q}_t = \{Q \in \mathcal{M}_{1,c} : Q = P|_{\mathcal{F}_t}\}$.

This finding is in contrast to the case of time-independent levels, where AV@R is in general not acceptance consistent. Another well-known example of a risk measure with time-dependent risk aversion parameters, where dynamic consistency has been fully investigated, is the dynamic entropic risk measure $(\rho_t^{ent})_{t\in I}$. It is defined by

$$\rho_t^{ent}(X) = \frac{1}{\gamma_t} \ln\left(\mathbb{E}\left[exp(-\gamma_t X) \mid \mathcal{F}_t\right]\right), \quad t \in I,\ X \in \mathcal{X},$$

with random risk aversion parameters $\gamma_t > 0$, $\gamma_t, \frac{1}{\gamma_t} \in L^\infty(\mathcal{F}_t)$ for all $t \in I$. As shown in [1, Proposition 1.43], the process $(\gamma_t)_{t\in I}$ determines time consistency properties of $(\rho_t^{ent})_{t\in I}$ as follows:

6.18 Proposition. *For $(\rho_t^{ent})_{t\in I}$ the following assertions hold:*

(i) $(\rho_t^{ent})_{t\in I}$ is rejection consistent if $\gamma_t \ge \gamma_{t+1}$ for all $t \in I$, $t < T$,

(ii) $(\rho_t^{ent})_{t\in I}$ is acceptance consistent if $\gamma_t \le \gamma_{t+1}$ for all $t \in I$, $t < T$,

(iii) $(\rho_t^{ent})_{t\in I}$ is time consistent if $\gamma_t = \gamma_0$ for all $t \in I$.

6.19 Remark. Although [1, Proposition 1.43] was proven for a slightly different definition of rejection sets and for a stronger version of rejection consistency, the result can also be applied to our theoretical framework.

While Theorems 6.13 and 6.15 are formulated for dynamic risk measures with constant risk aversion parameters, they do not cover the case of dynamic risk measures with time-dependent risk aversion parameters. For example, due to Theorem 6.13, the dynamic entropic risk measure with constant risk aversion parameters is both acceptance and rejection consistent; see [136, Example 2.4.8]. In order to add the case of time-dependent risk aversion parameters to Weber's results, we alter the definition of locally measure convexity.

6.20 Definition. For any $t \in I$ let $\mathcal{N}_t$ ($\mathcal{N}_t^c$) be the acceptance (rejection) set of the static risk measures $\psi_{(t)}$, where subindex (t) only labels the time-dependence of the parametrization of $\psi_{(t)}$ over time, i.e.,

$$\mathcal{N}_t = \{\mu \in \mathcal{M}_{1,c} : \psi_{(t)}(\mu) \le 0\}, \quad \mathcal{N}_t^c = \{\mu \in \mathcal{M}_{1,c} : \psi_{(t)}(\mu) > 0\}.$$

We say that the sequence $(\mathcal{N}_t)_{t\in I}$ of acceptance sets is *locally recursively measure convex* if for all $a \in \mathbb{R}$ and any ι on $\mathcal{N}_t \cap \mathcal{M}_1([-a,a])$, the mixture $\int_{\mathcal{N}_t} \nu\iota(d\nu)$ is an element of $\mathcal{N}_{t-1}$ for all $t \in I$. Analogously, $(\mathcal{N}_t^c)_{t\in I}$ is locally recursively measure convex if for all $a \in \mathbb{R}$ and any ι on $\mathcal{N}_t^c \cap \mathcal{M}_1([-a,a])$, the mixture $\int_{\mathcal{N}_t} \nu\iota(d\nu)$ is an element of $\mathcal{N}_{t-1}$ for all $t \in I$.

We then get the following result:

6.21 Theorem. *Let $(\psi_{(t)})_{t\in I}$ be a sequence of static risk measures, $\mathcal{N}_t \subseteq \mathcal{M}_{1,c}$ be its acceptance sets, $\mathcal{N}_t^c \subseteq \mathcal{M}_{1,c}$ its rejection sets, and $(\rho_t)_{t\in I}$ be the law-invariant dynamic risk measure on random variables defined by*

$$\rho_t(X) = \beta_t^T \cdot \psi_{(t)}\left(\mathcal{L}(X \mid \mathcal{F}_t)\right) \quad P\text{-a.s.} \tag{DynRM(t)}$$

If $(\mathcal{N}_t)_{t\in I}$ is locally recursively measure convex, then $(\rho_t)_{t\in I}$ is acceptance consistent, if $(\mathcal{N}_t^c)_{t\in I}$ is locally recursively measure convex, then $(\rho_t)_{t\in I}$ is rejection consistent.

Proof. This proof is essentially identical to the proof of [137, Theorem 4.2]. We only prove the case that $(\rho_t)_{t\in I}$ is acceptance consistent if $\mathcal{N}_t$ is locally recursively measure convex for any $t \in I$. Rejection consistency for locally recursively measure convex $\mathcal{N}_t^c$ can be proven similarly.

Let $t \in \{1,2,...,T-1\}$, $X \in \mathcal{X}$ and $a \in \mathbb{R}$ such that $X \in [-a,a]$. We define

$$\mathcal{Q}_t : \ (\Omega,\mathcal{F}_t) \to (\Omega,\mathcal{F}),\ \mathcal{Q}_t(\omega,A) = P(A \mid \mathcal{F}_t)(\omega) \quad \text{for } A \subseteq \Omega.$$

We set

$$\mu_s := \mathcal{L}(X \mid \mathcal{F}_s) \quad \text{for } s = t \text{ or } s = t+1.$$

Then, we obtain by disintegration for P-almost any $\omega \in \Omega$ that

$$\mu_t(\omega,\cdot) = \int \mu_{t+1}(\bar\omega,\cdot)\mathcal{Q}_t(\omega,d\bar\omega).$$

Suppose that $\rho_{t+1}(X) \le 0$. Then, by definition, $\psi_{(t+1)}(\mu_{t+1}) \le 0$, thus

$$\mu_{t+1}(\bar\omega,\cdot) \in \mathcal{N}_{t+1} \cap \mathcal{M}_1\left([-a,a]\right) \quad \text{for } P\text{-almost all } \bar\omega \in \Omega.$$

Hence, since $\mathcal{N}_t$ is locally recursively measure convex for any $t \in I$, for P-almost all $\omega \in \Omega$

$$\int \mu_{t+1}(\bar\omega,\cdot)P(d\bar\omega \mid \mathcal{F}_t)(\omega) \in \mathcal{N}_t,$$

thus

$$\mu_t \in \mathcal{N}_t \quad \text{for } P\text{-almost all } \omega \in \Omega.$$

This implies $\rho_t(X) \le 0$. Therefore $(\rho_t)_{t\in I}$ is acceptance consistent. □

Thus, if $(\mathcal{N}_t)_{t\in I}$ is locally recursively measure convex, which in simple terms means that $\mathcal{N}_t$ is getting "smaller" or that the parametrization of $(\psi_{(t)})_t$ reflects risk aversion that is increasing in time, then $(\rho_t)_{t\in I}$ as defined by (DynRM(t)) is acceptance consistent.

In order to apply the previous theorem to utility-based shortfall risk measures, we consider the "conditionally robust version":

$$\psi_{(t)}^{\text{rob SR}}(\mu) = \inf\{\eta \in \mathbb{R} : \mathbb{E}[u_t(X+\eta)] \ge u_t(0)\}, \text{ for } X \sim \mu,$$

as used in [1] and [132]; note that here and in the remainder of this section we consider increasing, concave functions $u_t : \mathbb{R} \to \mathbb{R}$ that are not identically constant. Consistency of conditionally robust utility-based shortfall risk measures is unaffected by shifts in u_t since then also the respective level $\lambda_t = u_t(0)$, see Example 3.1 is shifted. Yet, for this reason $\psi^{\text{rob SR}}$ is restrictive in the sense that the level λ cannot be chosen independently of u_t.

Another way to prevent inconsistency is to assume that all functions assume the same value in at least one point, that is $u_t(x_0) = u_s(x_0)$ for some $x_0 \in \mathbb{R}$ for any u_t, u_s under consideration. We denote utility-based shortfall risk measures as defined in Example 3.1, which satisfy this property, by $\psi^{\text{nor SR}}$.

6.22 Proposition. *If u_t is more concave than u_s, i.e., if there exists a strictly increasing concave function g such that $u_t = g \circ u_s$, then*

$$\psi_{(t)}^{nor\ SR}(\mu) \geq \psi_s^{nor\ SR}(\mu) \quad \textit{for all } \mu \in \mathcal{M}_{1,c}.$$

Proof. Since u_t and u_s are concave, increasing and $u_t(x_0) = u_s(x_0)$ for some $x_0 \in \mathbb{R}$, from u_t being more concave than u_s it follows that $u_t(x) \leq u_s(x)$ for all $x \in \mathbb{R}$. Consequently, we have that $\mathbb{E}[u_t(X)] \leq \mathbb{E}[u_s(X)]$ for all $X \in \mathcal{X}$ and therefore

$$\inf\{\eta \in \mathbb{R} : \mathbb{E}[u_t(X+\eta)] \geq \lambda\} \geq \inf\{\eta \in \mathbb{R} : \mathbb{E}[u_s(X+\eta)] \geq \lambda\}$$

for any $\lambda \in \mathbb{R}$ which concludes the proof. □

A similar result for optimal expected utility risk measures can be found in Proposition 4.11. We denote by $\rho_t^{\text{rob SR}}$ and $\rho_t^{\text{nor SR}}$ the dynamic risk measures which are defined by (DynRM(t)) and $\psi_{(t)}^{\text{rob SR}}$ or $\psi_{(t)}^{\text{nor SR}}$, respectively.

6.23 Corollary. *For $u \in \mathcal{U}$ it holds that*

(i) $(\rho_t^{rob/nor\ SR})_{t\in I}$ is acceptance consistent if $\varrho_{u_t}(x) \leq \varrho_{u_{t+1}}(x)$ for all $t \in I$ and all $x \in \mathbb{R}$.

(ii) $(\rho_t^{rob/nor\ SR})_{t\in I}$ is rejection consistent if $\varrho_{u_t}(x) \geq \varrho_{u_{t+1}}(x)$ for all $t \in I$ and all $x \in \mathbb{R}$.

Proof. The proof for $\rho_t^{\text{rob SR}}$ follows from [132, Corollary 5.4].

We show that $\varrho_{u_t}(x) \leq \varrho_{u_{t+1}}(x)$ implies acceptance consistency for $\rho_t^{\text{nor SR}}$. The case of rejection consistency works analogously.

From Proposition 3.8 we know that if $\varrho_{u_t}(x) \leq \varrho_{u_{t+1}}(x)$ for all $x \in \mathbb{R}$, then $u_t = g \circ u_{t+1}$ for a strictly increasing concave function g. Thus it follows from Proposition 6.22 that

$$\psi_{(t)}^{\text{nor SR}}(\mu) \leq \psi_{(t+1)}^{\text{nor SR}}(\mu) \quad \text{for all } \mu \in \mathcal{M}_{1,c}.$$

As this applies to all $t \in I$, we conclude that $(\rho_t^{\text{nor SR}})_{t\in I}$ is acceptance consistent. □

We exemplarily illustrate the results from this section for the consistency properties of the dynamic entropic risk measure as outlined in Proposition 6.18. To this end note that $(\rho_t^{ent})_{t\in I}$ is a special case of the utility-based shortfall risk measure for

$$u_t(x) = \frac{1}{\gamma_t}(1 - \exp(-\gamma_t x)).$$

If we assume that

$$\gamma_t \geq \gamma_{t+1} \quad \text{for all } t \in I,\ t < T,$$

then, due to Proposition 6.22, the sequence of the rejection sets

$$\mathcal{N}_t^c = \{\mu \in \mathcal{M}_{1,c} : \psi_{(t)}^{\text{nor SR}}(\mu) > 0\}$$

is locally recursively measure convex, thus $(\rho_t^{ent})_{t \in I}$ is rejection consistent due to Theorem 6.21. Moreover, for exponential utility we know that

$$\gamma_t = \varrho_{u_t} \quad \text{for all } t \in I,\ t < T,$$

thus rejection consistency of $(\rho_t^{ent})_{t \in I}$ follows immediately from Corollary 6.23. The case of acceptance consistency can be derived similarly.

CHAPTER 7

Recursively composed risk measures

In accordance with the literature and for the sake of a simpler notation, we consider discounted financial payoffs $X \in L^2(\Omega,\mathcal{F},P)$ and a finite time horizon $T > 0$ in this chapter. On the one hand we then choose a discrete-time approach to measure risk by recursively recalculating conditional risk on a small time grid $[t_k,t_{k+1})$ in order to finally assign a value to the overall recursive risk of a financial position X at finitely many time steps $t \in I = \{t_0 = 0, t_1, ..., t_N = T\}$. On the other hand we follow a continuous-time approach where we take the solution of a backward stochastic differential equation (BSDE) as the recursive risk of X at any time $t \in I = [0,T]$. Our goal is to bring these approaches together, and to study convergence from the discrete-time approach to the continuous-time approach.

Rosazza Gianian has shown in [121, Proposition 20] that a dynamic risk measure $(\rho_t)_{t\in[0,T]}$ can be identified as a conditional g-expectation and thus is the solution of a BSDE if $(\rho_t)_{t\in[0,T]}$ is a strictly monotone time consistent dynamic convex risk measure in a Brownian setting and if ρ_0 satisfies a certain boundedness condition. We have seen in Chapter 6 that dynamic risk measures such as V@R, AV@R and OEU are not time consistent in general. One can, however, construct time consistent dynamic risk measures in discrete time (DDRMs) by composing rescaled one-period risk measures by (ComRM). This relation allows us to compare composed risk measures with corresponding g-expectations which represents a major advantage for interpreting dynamic risk measures: The characteristics of g are related to properties of the respective dynamic risk measure and the functional g which could also depend on preferences of the investor and on parameters linked to the evaluated financial position, is easy to interpret, see [121]; the risk evaluation by composed DDRMs, on the other hand, is well understood in any single time interval.

The main contribution of this chapter is the implementation of composed scaled risk measures as proposed by Stadje [126] and the distinction between this approach and the concept of Romanovski [120]. We provide a theoretical explanation for our new numerical results on the convergence behavior of the different approaches and point out that our results indicate that risk measures can be divided into two groups: Whereas risk measures of the form

$$\rho^{\lambda}(X) = \mathbb{E}[-X] + b(\lambda)\sigma(X),$$

need additional scaling when recursively composed, risk measures of the form

$$\rho^{\lambda}(X) = \mathbb{E}[-X] + b(\lambda)\sigma^2(X),$$

are perfectly suited for convergence from a discrete time setting to a continuous time setting without further scaling; where $\sigma(X)$ is the standard deviation of X and $b : \mathbb{R}_+ \to \mathbb{R}$ includes the predetermined parameter $\lambda \in \mathbb{R}$. We start this chapter with an exemplary presentation of the convergence behavior of certainty equivalents in Section 7.1 where we encounter the exact same two groups of functionals with respect to convergence in time. Furthermore, this chapter contains the following sections: In 7.2 we present a first naive approach of composed risk measures, an alternative scaled composition of risk measures is shown in 7.3 and we illustrate the convergence properties of both approaches in 7.4.

7.1 Two classes of Arrow-Pratt approximation for certainty equivalents

The common definition of the *Arrow-Pratt approximation* (APA) is inseparably linked to the concept of certainty equivalents. Following [81, Section 3.4], we now consider a random variable $Y = k(\mu + X)$, with $\mathbb{E}[X] = 0$ (thus X is called *pure risk*) and let $g(k)$ denote its associated certainty equivalent $C_u(k(\mu + X))$, that is the sure amount that makes an investor indifferent between investing in Y, or receiving the sure gain $C_u(Y)$, i.e., it verifies:

$$\mathbb{E}\left[u(k(\mu + X))\right] = u(g(k))$$

for any utility function u. APA helps us to understand the characteristics of certainty equivalents for small risks. We make use of this concept when it comes to examining the behavior of utility on time grids of decreasing length; here, that is $k \to 0$. Clearly, we observe $g(0) = 0$, and we get

$$\mathbb{E}\left[(\mu + X)u'\left(k(\mu + X)\right)\right] = g'(k)u'\left(g(k)\right)$$

so that $g'(0) = \mu$, and

$$\mathbb{E}\left[(\mu + X)^2 u''\left(k(\mu + X)\right)\right] = \left(g'(k)\right)^2 u''\left(g(k)\right) + g''(k)u'\left(g(k)\right),$$

which implies

$$g''(0) = \frac{u''(0)}{u'(0)}\mathbb{E}[X^2].$$

Using a Taylor expansion of g around $k = 0$, we obtain that

$$C_u(k(\mu + X)) \simeq k\mu - \frac{1}{2}k^2\varrho(0)\mathbb{E}[X^2],$$

where $\varrho(z)$ is the absolute risk aversion as defined in Section 3.2. This is the APA for certainty equivalents. It states that the certainty equivalent for a small pure risk is approximately proportional to its variance. As the size k of this risk tends to zero, its certainty equivalent tends to zero by k^2.

7.1 Example. Let us consider certainty equivalents of a financial position (X_t) and let us restrict our view on a time step $t \to t+h$, i.e., we define C_u on $[t,t+h]$ as follows:

$$C_{u,t}: \ L^2(\mathcal{F}_{t+h}) \to L^2(\mathcal{F}_t): \ C_{u,t}(X_{t+h}) := u^{-1}\left(\mathbb{E}_t\left[u(X_{t+h})\right]\right), \ \text{ where } \mathbb{E}_t[X] := \mathbb{E}[X \mid \mathcal{F}_t].$$

The Arrow-Pratt approximation is then:

$$\begin{aligned}
C_{u,t}(X_{t+h}) &= u^{-1}\left(\mathbb{E}_t\left[u(X_{t+h})\right]\right)\\
&= u^{-1}\Bigg(\mathbb{E}_t\Bigg[u(\mathbb{E}_t[X_{t+h}]) + u'(\mathbb{E}_t[X_{t+h}])(X_{t+h} - \mathbb{E}_t[X_{t+h}])\\
&\qquad\qquad + \frac{1}{2}u''(\mathbb{E}_t[X_{t+h}])(X_{t+h} - \mathbb{E}_t[X_{t+h}])^2\Bigg]\Bigg)\\
&= u^{-1}\Bigg(u(\mathbb{E}_t[X_{t+h}]) + u'(\mathbb{E}_t[X_{t+h}])\underbrace{\mathbb{E}_t[X_{t+h} - \mathbb{E}_t[X_{t+h}]]}_{=0}\\
&\qquad\qquad + \frac{1}{2}u''(\mathbb{E}_t[X_{t+h}])\underbrace{\mathbb{E}_t\left[(X_{t+h} - \mathbb{E}_t[X_{t+h}])^2\right]}_{=Var_t[X_{t+h}]}\Bigg)\\
&= \mathbb{E}_t[X_{t+h}] + \frac{1}{u'(\mathbb{E}_t[X_{t+h}])}\left(\frac{1}{2}u''(\mathbb{E}_t[X_{t+h}])Var_t[X_{t+h}]\right)\\
&\qquad - \frac{1}{2}\frac{u''(\mathbb{E}_t[X_{t+h}])}{(u'(\mathbb{E}_t[X_{t+h}]))^3}\left(\frac{1}{2}u''(\mathbb{E}_t[X_{t+h}])Var_t[X_{t+h}]\right)^2.
\end{aligned}$$

If we take X as a geometric Brownian motion

$$dX_t = \mu dt + \sigma dW_t,$$

this approximation takes the form:

$$C_{u,t}(X_{t+h}) = \mu h + \frac{1}{2}\left(\frac{u''}{u'}\right)(\mu h)\sigma^2 h + o(h^2).$$

Thus, for exponential utility $u(x) = \frac{1}{\gamma}(1 - \exp(-\gamma x))$, the APA of the certainty equivalent on $[t,t+h]$ is:

$$C_{u,t}(X_{t+h}) = \mu h - \frac{1}{2}\gamma\sigma^2 h + o(h^2).$$

We notice that in the previous example $C_{u,t}(X_{t+h})$ converges to $C_{u,t}(X_t)$ by the factor h. But there is another class of certainty equivalents for which speed of convergence for $h \to 0$ is $\sqrt{h}$:

7.2 Example. We consider the λ-*quantile CE*, which is a certainty equivalent in the sense of Kreps and Porteus [96], on a small time step $[t,t+h]$ defined as:

$$\mathfrak{m}_t^{(\lambda)}: \ L^2(\mathcal{F}_{t+h}) \to L^2(\mathcal{F}_t): \ \mathfrak{m}_t(X_{t+h}) := \text{ess sup}\{m_t \in L^2(\mathcal{F}_t) \ : \ P(-X_{t+h}+m_t < 0 \mid \mathcal{F}_t) \le \lambda\}.$$

If we assume the financial position (X_t) to be a geometric Brownian motion $dX_t = \mu dt + \sigma dW_t$, i.e., $X_{t+h} \sim \mathcal{N}(\mu h, \sigma^2 h)$, then, as shown in Remark A.5, the λ-quantile CE takes the form

$$\mathfrak{m}_t^{(\lambda)}(X_{t+h}) = \mu h + \sigma\sqrt{h}\Phi^{-1}(1-\lambda).$$

Comparing $C_{u,t}(X_{t+h})$ from Example 7.1 to $\mathfrak{m}_t^{(\lambda)}(X_{t+h})$ from Example 7.2, we observe different convergence behavior for $h \to 0$. This may surprise readers who are familiar with convergence results for recursive utility. In the upcoming sections we point out how this finding strongly relates to the results from [126] for dynamic risk measures.

7.2 Composed risk measures and g–expectations

We consider the following stochastic differential equations for the continuous time setting of risk evaluation: Take the backward stochastic differential equation

$$\begin{aligned} -dY_t &= g(t,Y_t,Z_t)dt - Z_t dW_t, \\ Y_T &= X, \end{aligned} \tag{BSDE}$$

for all $t \in [0,T]$, where we refer to Section 2.5 for technical assumptions on g that ensure a unique solution $(Y_t,Z_t)_{t\in I}$, and the forward stochastic differential equation

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma S_t dW_t, \\ S_0 &= s, \end{aligned} \tag{FSDE}$$

for all $t \in [0,T]$, that is a Black-Scholes model which implies that logarithmic returns follow a normal distribution with expectation μ and variance σ^2. Together we call this a *forward backward stochastic differential equation*:

$$\begin{aligned} S_0 &= s, \\ dS_t &= \mu S_t dt + \sigma S_t dW_t, \\ Y_T &= \Phi(S_T), \\ -dY_t &= g(t,Y_t,Z_t)dt - Z_t dW_t, \end{aligned} \tag{FBSDE}$$

for all $t \in [0,T]$.

7.2.1 Discrete-time approximation of BSDEs

In the following W is a standard Brownian motion and μ, σ, g are valued respectively in $\mathbb{R}^n$ and $\mathbb{R}$. We know from Section 2.5 that we can construct time consistent risk measures by choosing a suitable driver g for (BSDE). In order to solve (FBSDE), we have to numerically approximate the solution. For this concern, given a discrete time grid $I = \{t_k \in [0,T] : t_k = kh = k\frac{T}{N},\ 0 \le k \le N\}$, we consider a discretization of (BSDE) as presented in Bouchard and Touzi [23]:

$$Y^N_{t_{k+1}} - Y^N_{t_k} = -g(t_k, Y^N_{t_k}, Z^N_{t_k})h + Z^N_{t_k}\Delta W_{k+1}, \tag{7.1}$$

where we use the notation $\Delta W_{k+1} := W^N_{t_{k+1}} - W^N_{t_k}$. (FSDE) leads to the final data $\tilde{Y}_T = \Phi(S_T)$, and, in order to approximate the forward component, we use a standard Euler scheme

$$S^N_{t_{k+1}} - S^N_{t_k} = \mu(t_k, S^N_{t_k})h + \sigma(t_k, S^N_{t_k})\Delta W_{k+1}.$$

This leads to $Y^N_{t_N} = \Phi(S^N_{t_N})$ which is needed in addition to (7.1) for the discretization of (BSDE). Since, given $(Y_{t_{k+1}}, Z_{t_{k+1}})$, there are no $\mathcal{F}_{t_k}$-measurable random variables (Y_{t_k}, Z_{t_k}) which satisfy (7.1), we refer to the following backward procedure for the definition of the discrete-time approximation $(Y^N_t, Z^N_t)_{t\in I}$ which was given in [23]:

$$\begin{aligned} Y^N_{t_N} &= \Phi(S_T), \\ Z^N_{t_N} &= 0, \\ Z^N_{t_k} &= \frac{1}{h}\mathbb{E}\left[Y^N_{t_{k+1}}\Delta W_{k+1} \mid \mathcal{F}_{t_k}\right], \\ Y^N_{t_k} &= \mathbb{E}\left[Y^N_{t_{k+1}} \mid \mathcal{F}_{t_k}\right] + hg(t_k, Y^N_{t_k}, Z^N_{t_k}). \end{aligned} \tag{7.2}$$

The equality for $Z^N_{t_k}$ was acquired by multiplying (7.1) by ΔW_{k+1} and taking conditional expectation with respect to $\mathcal{F}_{t_k}$. The representation of $Y^N_{t_k}$ follows from (7.1) by taking expectation with respect to $\mathcal{F}_{t_k}$.

Briand, Delyon and Mémin [26, Theorem 2.1] show that under certain assumptions $(Y^N_{t_k}, Z^N_{t_k})$ from (7.1), which is based on M-dimensional random walks W^N converging to the M-dimensional Brownian motion W underlying (BSDE), converge to the solution (Y, Z) of (BSDE). This result is the foundation of the numerical implementation in Section 7.4.

7.2.2 A recursive construction of time consistent risk measures

At this point we want to introduce a construction of a time consistent DDRM $(\rho^{\text{Com}}_t)_{t\in I}$ from an arbitrary DDRM $(\rho_t)_{t\in I}$. The recursive construction of the *composed risk measure* (ComRM) is defined as

$$\begin{aligned} \rho^{\text{Com}}_{t_N}(X) &:= \rho_{t_N}(X) = -X \\ \rho^{\text{Com}}_{t_k}(X) &:= \rho_{t_k,t_{k+1}}(-\rho^{\text{Com}}_{t_{k+1}}(X)), \quad t_k \le t_{k_{N-1}}, \end{aligned} \tag{ComRM}$$

where $\rho_{t_k,t_{k+1}}$ is the restriction of ρ_{t_k} to $L^2(\mathcal{F}_{t_{k+1}})$.

7.2.3 Composed value at risk

Generally, for time-independent levels $\lambda_t := \lambda$, $(V@R^\lambda_t)_{t\in I}$ is not time consistent as shown in Cheridito and Stadje [36, Example 3.1] and in [69, Example 11.13]. However, $(V@R^\lambda_t)_{t\in I}$ is acceptance and rejection consistent; see [136, Example 2.4.7]. In this section we consider the time consistent composed value at risk in the sense of (ComRM):

$$\begin{aligned} ComV@R^\lambda_{t_N}(X) &:= V@R^\lambda_{t_N}(X) = -X \\ ComV@R^\lambda_{t_k}(X) &:= V@R^\lambda_{t_k,t_{k+1}}(-ComV@R^\lambda_{t_{k+1}}(X)), \quad t_k \le t_{k_{N-1}}. \end{aligned} \tag{ComV@R}$$

Suppose we wish for (7.1) to represent ComV@R. Under this numerical approximation $Y^N_{t_{k+1}}$ conditional on $\mathcal{F}_{t_k}$ is normally distributed with mean $Y^N_{t_k} - g(t_k, Y^N_{t_k}, Z^N_{t_k})h$ and standard deviation $\sqrt{h}\,|Z^N_{t_k}|$. This fact implicates some nice properties of ComV@R as shown in the following remarks.

7.3 Remark. (i) $V@R$ corresponding to a normally distributed financial position X with variance $\sigma^2(X)$ is

$$V@R^{\lambda}(X) = \mathbb{E}[-X] + \sigma(X)\Phi^{-1}(1-\lambda),$$

where Φ^{-1} is the inverse cumulative distribution function of the standard normal distribution; see Remark A.5.

(ii) $V@R$ is subadditive for normally distributed financial positions if $\lambda \le 0.5$; see Embrechts, McNeil and Straumann [56, Theorem 1], and Remark A.6.

(iii) Typically, return rates of well diversified portfolios of profit and loss distributions of large companies are not very different from normal, as pointed out in the analysis of the S&P 500 in [55, pages 11-12 and page 29].

The following remark is to justify why ComV@R is assumed to be nearly coherent if we consider a large number N of time steps.

7.4 Remark. 1. The price dynamics of the underlying asset are given by a geometric Brownian motion with drift:

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma S_t dW_t, \\ S_0 &= s, \end{aligned}$$

where W_t is a P-Brownian motion.

2. S_t is log-normally distributed with parameters

$$\mathbb{E}[S_t] = s\exp(\mu t),$$

and

$$Var[S_t] = s^2\exp(2\mu t)\left(\exp(\sigma^2 t) - 1\right).$$

Therefore, in small time steps h, S changes by an amount that is "nearly normally" distributed with mean $\tilde{\mu}Sh$, $\tilde{\mu} := \mu - \frac{1}{2}\sigma^2$ and standard deviation $\sigma S\sqrt{h}$. Thus the loss involved in the asset is locally "nearly normal" on a small time grid; see Remark A.7.

3. V@R is subadditive for normally distributed random variables, see Remark 7.3 (ii), and therefore V@R of the loss involved in the asset is locally "nearly coherent".

4. (ComRM) inherits subadditivity of V@R, thus also ComV@R is "nearly coherent" if we consider a large number of time steps.

We set $Y^N_{t_N} = -X$. By discretization (7.1), we then get

$$ComV@R^N_{t_N}(Y^N_{t_N}) = Y^N_{t_N},$$

and

$$\begin{aligned} ComV@R^N_{t_{N-1}}(Y^N_{t_N}) &= V@R^N_{t_{N-1},t_N}(-ComV@R^N_{t_N}(Y^N_{t_N})) \\ &= \mathbb{E}\left[Y^N_{t_N} \mid \mathcal{F}_{t_{N-1}}\right] + \Phi^{-1}(1-\lambda)\left|Z^N_{t_{N-1}}\right|\sqrt{h}. \end{aligned} \tag{7.3}$$

Comparing (7.3) and (7.2), we conclude that by choosing the driver

$$g^{\mathrm{V@R}}(t,y,z) = \frac{\Phi^{-1}(1-\lambda)}{\sqrt{h}}\,|z|$$

we should get a dynamic risk measure in continuous time (CDRM) corresponding to ComV@R. But, due to [126, Proposition 5.1], ComV@R with given constant level λ does not generally converge to a BSDE solution for $N \to \infty$ if it is not properly rescaled. This is intuitively obvious because of the form of $g^{\mathrm{V@R}}(t,y,z)$: If we insert $g^{\mathrm{V@R}}$ in (7.2), it can easily be seen that:

$$g^{\mathrm{V@R}}(t,y,z) \to \infty \quad \text{as } N \to \infty.$$

We follow up on this in Section 7.3 where we consider composed scaled risk measures as a robust alternative to (ComRM).

7.2.4 Composed average value at risk

Due to [69, Example 11.13], $(AV@R^{\lambda_t}_t)_{t\in I}$ is not time consistent for time-independent levels $\lambda_t = \lambda$ and we therefore rather consider the composed average value at risk in the sense of (ComRM) which inherits the coherency property from all the AV@Rs used for the composition. Note again that discretization (7.1) leads to normally distributed $Y^N_{t_k}$. As shown in Remark A.8, for normally distributed financial position X with variance $\sigma^2(X)$ we have:

$$AV@R^{\lambda}(X) = \mathbb{E}[-X] + \frac{\sigma(X)}{\lambda\sqrt{2\pi}}\exp\left(-\frac{\Phi^{-1}(1-\lambda)^2}{2}\right),$$

which means that we get

$$ComAV@R^N_{t_{N-1}}(Y^N_{t_N}) = \mathbb{E}\left[Y^N_{t_N} \mid \mathcal{F}_{t_{N-1}}\right] + \left|Z^N_{t_{N-1}}\right|\sqrt{h}\frac{1}{\lambda\sqrt{2\pi}}\exp\left(-\frac{\Phi^{-1}(1-\lambda)^2}{2}\right). \tag{7.4}$$

Comparing (7.4) and (7.2), we conclude that by choosing the driver

$$g^{\mathrm{AV@R}}(t,y,z) = \frac{1}{\lambda\sqrt{2\pi h}}\exp\left(-\frac{\Phi^{-1}(1-\lambda)^2}{2}\right)|z|$$

we get a CDRM corresponding to ComAV@R. However, also ComAV@R with given constant level λ does not generally converge to a BSDE solution for $N \to \infty$ if it is not rescaled, and

$$g^{\mathrm{AV@R}}(t,y,z) \to \infty \quad \text{as } N \to \infty.$$

Instead, we suggest to choose ComScaAV@R as introduced in Section 7.3.

7.2.5 Composed entropic risk measure

The entropic risk measure for normally distributed X with variance $\sigma^2(X)$ is

$$\rho^{\text{ent}}(X) = -\mathbb{E}[X] + \frac{1}{2}\gamma\sigma^2(X).$$

Note that this coincides with the APA of the certainty equivalent from Example 7.1 for the exponential utility function $u(x) = \frac{1}{\gamma}(1 - \exp(-\gamma x))$ with time-independent parameter $\gamma_t = \gamma > 0$. If we wish for (7.1) to represent ComENT, we get:

$$ComENT^N_{t_{N-1}}(Y^N_{t_N}) = \mathbb{E}[Y^N_{t_N} \mid \mathcal{F}_{t_{N-1}}] + \frac{1}{2}\gamma Var[Y^N_{t_N} \mid \mathcal{F}_{t_{N-1}}] \tag{7.5}$$

and the comparison of (7.5) and (7.2) leads to the driver

$$g^{\text{ENT}}(t,y,z) = \frac{1}{2}\gamma z^2.$$

This is in line with Barrieu and El Karoui [13, Proposition 6.4] which, for the sake of completeness, we state again at this point.

7.5 Proposition (Proposition 6.4 from [13]). *The dynamic entropic risk measure $(\rho_t^{ent})_{t\in I}$ with time-independent parameter $\gamma_t = \gamma > 0$ is solution of the following BSDE with the quadratic coefficient $g^{ENT}(t,z) = \frac{1}{2}\gamma z^2$ and terminal bounded condition:*

$$-d\rho_t^{ent}(X) = \frac{1}{2}\gamma Z_t^2 dt - Z_t dW_t,$$
$$\rho_T^{ent}(X) = -X.$$

We note that the entropic risk measure for normally distributed random variables is of the form

$$\rho^\lambda(X) = \mathbb{E}[-X] + b(\lambda)\sigma^2(X),$$

and that no constraint on g or additional scaling of the composed risk measure is necessary to ensure that $(ComENT_t^N)_{t\in I}$ converges to a continuous-time risk measure $(\rho_t^{\text{ent},g})_{t\in[0,T]}$. We further note that the known fact that ρ^{ent} is a more conservative risk measure the bigger the parameter γ can also be seen from the form of g: The bigger γ, the bigger g and thus the more conservative $\rho^{\text{ent},g}$; see Section 2.5.

For a fixed risk aversion parameter $\gamma > 0$, the *dynamic entropic risk measure* is time consistent; see Remark A.4.

7.3 Composed scaled risk measures and g–expectations

Stadje [126, Proposition 5.1] shows that under certain conditions all one-period coherent risk measures explode in the limit of the number of time steps N if they are not properly rescaled. This happens for the following reason: First, note that the standard deviation of the increments of the Brownian motion which is used to model X is of order $\sqrt{h}$. Since V@R and AV@R are risk measures of the form

$$\rho^\lambda(X) = \mathbb{E}[-X] + b(\lambda)\sigma(X),$$

as the size of the time steps goes to zero by the factor h, the involved risk X in the according time interval goes to zero by the factor $\sqrt{h}$. Intuitively, one could say that in our framework the risk of, say a loan default happens in $[t_k,t_{k+1}]$ decreases by a factor $\sqrt{h}$ which is smaller than the decrease of time when we consider increasingly small time intervals. Consequently, BSDE-drivers g of order $\frac{1}{\sqrt{h}}$ do not converge to corresponding composed risk measures for $N \to \infty$, where N is the number of time steps; see Section 7.4 for numerical evidence of this. It is therefore necessary to introduce another construction of time consistent risk measures.

Following [126, page 14], we present a scaled construction of time consistent risk measures from arbitrary risk measures ρ_{t_k}. To this end we introduce scaled one-period risk measures $\rho_{t_k,t_{k+1}}$ by:

$$\tilde{\rho}_{t_k,t_{k+1}}(X) := \mathbb{E}[-X \mid \mathcal{F}_{t_k}] + h\ \rho_{t_k,t_{k+1}}\left(\frac{1}{\sqrt{h}}(X - \mathbb{E}[X \mid \mathcal{F}_{t_k}])\right)$$

for any $\mathcal{F}_{t_{k+1}}$-measurable X. Broadly speaking, $\tilde{\rho}_{t_k,t_{k+1}}$ is the sum of the conditional expected value of $-X$ and the original risk measure $\rho_{t_k,t_{k+1}}$ which is taken of the (variance-independent) value $\frac{1}{\sqrt{h}}(X - \mathbb{E}[X \mid \mathcal{F}_{t_k}])$ and is additionally multiplied by h to ensure convergence for $N \to \infty$. Note that, if ρ_{t_k} is positively homgeneous, $\rho_{t_k,t_{k+1}}$ takes the form:

$$\tilde{\rho}_{t_k,t_{k+1}}(X) = \mathbb{E}[-X \mid \mathcal{F}_{t_k}] + \sqrt{h}\ \rho_{t_k,t_{k+1}}\left(X - \mathbb{E}[X \mid \mathcal{F}_{t_k}]\right),$$

and, if $-X \sim \mathcal{N}(\mu,\sigma^2)$, then $-(X - \mathbb{E}[X]) \sim \mathcal{N}(0,\sigma^2)$. In this form $\tilde{\rho}_{t_k,t_{k+1}}(X)$ is the estimated value of the risk X at the end of the considered time interval given the information at the beginning of the interval plus the scaled risk of a deviation of the risk from its expected value.

Then, the recursive construction of the *composed scaled risk measure* (ComScaRM) is

$$\begin{aligned} \rho^{\text{ComSca}}_{t_N}(X) &:= \rho_{t_N}(X) = -X \\ \rho^{\text{ComSca}}_{t_k}(X) &:= \tilde{\rho}_{t_k,t_{k+1}}(-\rho^{\text{ComSca}}_{t_{k+1}}(X)), \quad t_k \le t_{k_{N-1}}. \end{aligned} \tag{ComScaRM}$$

7.6 Example. (a) If we want the discretized BSDE (7.1) to represent a composed version of V@R in the sense of (ComScaRM), we get:

$$\begin{aligned} ComScaV@R^N_{t_{N-1}}(Y^N_{t_N}) &= \mathbb{E}[Y^N_{t_N} \mid \mathcal{F}_{t_{N-1}}] \qquad (7.6)\\ &\quad + h\left(V@R_{t_{N-1},t_N}\left(\frac{1}{\sqrt{h}}\left(-Y^N_{t_N} - \mathbb{E}[-Y^N_{t_N} \mid \mathcal{F}_{t_{N-1}}]\right)\right)\right) \\ &= \mathbb{E}[Y^N_{t_N} \mid \mathcal{F}_{t_{N-1}}] + h\left(\frac{1}{\sqrt{h}}\left(\Phi^{-1}(1-\lambda)\left|Z^N_{t_{N-1}}\right|\sqrt{h}\right)\right) \\ &= \mathbb{E}[Y^N_{t_N} \mid \mathcal{F}_{t_{N-1}}] + \Phi^{-1}(1-\lambda)\left|Z^N_{t_{N-1}}\right| h, \qquad (7.7) \end{aligned}$$

where we assumed (nearly) coherency of ComScaV@R as justified for ComV@R and we made use of the fact that $-Y^N_{t_N} - \mathbb{E}[-Y^N_{t_N} \mid \mathcal{F}_{t_{N-1}}]$ conditional on $\mathcal{F}_{t_{N-1}}$ is normally distributed with mean 0 and standard deviation $\sqrt{h}\left|Z^N_{t_{N-1}}\right|$. The comparison of (7.6) and (7.2) leads to the driver

$$g^{\text{V@R}}_{\text{Stadje}}(t,y,z) = \Phi^{-1}(1-\lambda)\,|z|\,.$$

This is exactly what we hoped to get: A driver $g^{\text{V@R}}_{\text{Stadje}}$ which is independent of the size of the time step h and therefore can't "explode" as $h \to 0$. $g^{\text{V@R}}_{\text{Stadje}}$ can be viewed as the continuous-time analog of the discrete "driver" characterizing the one-period risk. It is actually a driver of the form $g(t,y,z) = a\,|z|$, where a is some constant. This is in line with the theory of discrete BSDEs: Any time consistent nonlinear expectation corresponds to a driver of this form; see Cohen and Elliott [38].

(b) Equivalently, if we want the discretized BSDE (7.1) to represent a composed version of AV@R in the sense of (ComScaRM), we get:

$$\begin{aligned} ComScaAV@R^N_{t_{N-1}}(Y^N_{t_N}) &= \mathbb{E}[Y^N_{t_N} \mid \mathcal{F}_{t_{N-1}}] \\ &\quad + \sqrt{h}AV@R_{t_{N-1},t_N}\left(-Y^N_{t_N} - \mathbb{E}[-Y^N_{t_N} \mid \mathcal{F}_{t_{N-1}}]\right) \\ &= \mathbb{E}[Y^N_{t_N} \mid \mathcal{F}_{t_{N-1}}] + \frac{1}{\lambda\sqrt{2\pi}}\exp\left(-\frac{\Phi^{-1}(1-\lambda)^2}{2}\right)\left|Z^N_{t_{N-1}}\right| h. \end{aligned} \tag{7.8}$$

The comparison of (7.8) and (7.2) then leads to the driver

$$g^{\text{AV@R}}_{\text{Stadje}}(t,y,z) = \frac{1}{\lambda\sqrt{2\pi}}\exp\left(-\frac{\Phi^{-1}(1-\lambda)^2}{2}\right)|z|,$$

which again is independent of h.

(c) We remark that scaling is not necessary for the dynamic entropic risk measure since the composition (ComRM) already converges to a continuous-time version of ρ^{ent}. At this point note that [126, Proposition 5.1] only includes coherent risk measures. We therefore assume that there exist two classes of risk measures in terms of convergence from a discrete-time composition to a continuous-time dynamic risk measure just like we see in Section 7.1 for certainty equivalents on small time grids.

Note that in the previous sections we have not found a "dynamic (A)V@R in continuous time", but rather figured out CDRMs which correspond to time consistent DDRMs which are composed by the initial static risk measures. Thus the initial risk measures and the obtained CDRMs are related, but since the construction of the time consistent risk measures is a backwards composition of one-period risk measures, the resulting risk measures are not directly comparable to the initial risk measures.

7.4 Numerical analysis of the convergence of composed (scaled) risk measures and g–expectations

The numerical approximation of the solutions to BSDEs in this section is based on [23] and [80]. In this context, we also refer to Douglas, Ma and Protter [52] and Ma, Protter and Yong [99]. Hereby, the simulation of the backward component Y is possible due to the following result: Under standard Lipschitz conditions, the pair (Y,Z), which solves the BSDE, can be expressed as a function of the forward process S, i.e., $(Y_t,Z_t) = (u(t,S_t),v(t,S_t))$, $t \le 1$, for some deterministic functions u and v. This is the general idea of the algorithm we use, see also Listings A.1, A.2, A.3:

1. Using a standard Euler approximation scheme, we generate M Monte Carlo paths of the forward process $S_{t_k}^N$ which are used in the further steps of the algorithm:

$$\begin{aligned} S_{t_{k+1}}^N - S_{t_k}^N &= \mu S_{t_k}^N h + \sigma S_{t_k}^N \Delta W_{k+1}, \\ S_{t_0}^N &= s. \end{aligned}$$

2. From [23, Remark 3.1] we know that

$$\mathbb{E}[Y_{t_{k+1}}^N \Delta W_{k+1} \mid \mathcal{F}_{t_k}] = \mathbb{E}[Y_{t_{k+1}}^N \Delta W_{k+1} \mid S_{t_k}^N]$$

   and

$$\mathbb{E}[Y_{t_{k+1}}^N \mid \mathcal{F}_{t_k}] = \mathbb{E}[Y_{t_{k+1}}^N \mid S_{t_k}^N].$$

   Therefore we can represent $(Y_{t_k}^N, Z_{t_k}^N)$ by a projection on a finite basis of functions:

$$Y_{t_k}^N = \alpha_k p_k, \quad Z_{t_k}^N = \beta_k p_k,$$

   where $p_k = p_k(S_{t_k}^N)$.

3. (BSDE) solver:
   Following [80], for every backwards step $t_{k+1} \to t_k$ we solve a least squares problem (3 Picard iterations) to attain α_k, β_k. This leads to an approximation of $Y_{t_k}^N$ and $Z_{t_k}^N$ which yields an approximate solution of the discretized (BSDE):

$$Y_{t_{k+1}}^N - Y_{t_k}^N = -g(t_k, Y_{t_k}^N, Z_{t_k}^N)h + Z_{t_k}^N \Delta W_{k+1}.$$

4. (ComRM), (ComScaRM) solver:
   For every backwards step we use a projection on the same finite basis α_k as in our (BSDE) solver to get an approximation of $Y_{t_k}^N$. Having the expected value of $Y_{t_k}^N$ the calculation is given by (ComRM) or (ComScaRM), respectively, e.g.,

$$ComScaV@R_{t_{N-1}}^N(Y_{t_N}^N) = \mathbb{E}[Y_{t_N}^N \mid \mathcal{F}_{t_{N-1}}] + \sqrt{h} V@R_{t_{N-1},t_N}\left(-Y_{t_N}^N - \mathbb{E}[-Y_{t_N}^N \mid \mathcal{F}_{t_{N-1}}]\right).$$

We refer to [80] for a more comprehensive description of the BSDE approximation and for results on error estimations due to the approximation. In the upcoming examples, which are inspired by Romanovski [120], we consider portfolios with the following payoffs:

1. Smooth payoff: $X := -\ln(S_T + 1)$,
2. Call: $X := -(S_T - 100)^+$,
3. Barrier payoff: $X := -70 \cdot \mathbf{1}_{\{S_T > 135\}}$.

The barrier payoff is either 0 or -70 in contrast to the other payoffs which react "more smoothly" on changes in S_T. We set $S_0 = 100$, $\mu = 0.06$, $\sigma = 0.2$, $T = 0.5$ and the level of (average) value at risk at $\lambda = 0.05$. To make optimal use of our computational resources and to minimize numerically caused fluctuations in the results, we take the following number of paths of the underlying:

Table 7.1: Number of paths.

number of time steps N	**paths of underlying M**
2 - 256	100,000
512	50,000

We use a Hermite polynomial base of degree 9 in order to get results that are best comparable to [120, Section 3.2]. In order to ensure repeatability of our results, we initialize MATLAB's random number generator using the seed 1. We then calculate CDRMs via g-expectations at $t = 0$ as:

$$\rho^g(X) := \mathcal{E}_g[-X] = Y_0^{-X},$$

and also present composed scaled DDRMs at $t = 0$:

$$\rho_0^{\text{ComRM}}(X) \quad \text{or} \quad \rho_0^{\text{ComScaRM}}(X).$$

Following the theoretical insights from the previous sections for value at risk and average value at risk, we consider how composed risk measures converge to risk measures from g-expectations for different portfolios.

As a first result, we consider the $1 - \lambda = 0.95$-quantiles of the M simulations for S_T for N time steps together with the (average) value at risks of the respective payoffs in Table 7.2.

Table 7.2: Quantiles and risk measures.

	S_T	Call		Barrier payoff		Smooth payoff	
N	0.95-quantile	V@R	AV@R	V@R	AV@R	V@R	AV@R
2	121.89	21.89	27.43	0.00	5.60	4.81	4.85
4	124.60	24.60	31.24	0.00	14.70	4.83	4.88
8	126.73	26.73	34.11	0.00	22.40	4.85	4.90
16	127.70	27.70	35.67	0.00	28.00	4.86	4.92
32	128.12	28.12	36.78	0.00	31.50	4.86	4.92
64	128.45	28.45	37.16	0.00	32.20	4.86	4.93
128	128.49	28.49	36.74	0.00	32.20	4.86	4.92
256	128.74	28.74	37.14	0.00	33.60	4.86	4.93
512	128.81	28.81	37.13	0.00	32.90	4.87	4.93

Table 7.2 shows that both, V@R and AV@R, converge to rather constant values for increasing N. Note that generally V@R<AV@R and that for the barrier payoff we have that V@R=0 whereas AV@R>0 which is intuitively obvious due to the definition of AV@R; see Example 2.4 (b).

7.4.1 Approximation results for composed risk measures and unscaled g–expectations

In this section we consider composed risk measures and g-expectations as presented in Section 7.2. The results give numerical evidence for the theoretical insight that the naively obtained BSDE drivers for value at risk

$$g^{\mathrm{V@R}}(t,y,z) = \frac{\Phi^{-1}(1-\lambda)}{\sqrt{h}}\,|z|$$

and average value at risk

$$g^{\mathrm{AV@R}}(t,y,z) = \frac{1}{\lambda\sqrt{2\pi h}}\exp\left(-\frac{\Phi^{-1}(1-\lambda)^2}{2}\right)|z|$$

do not imply CDRMs that converge to composed risk measures as $N \to \infty$; see Table 7.3.

Table 7.3: Composed risk measures and risk measures from g-expectations.

	Call		Barrier payoff		Smooth payoff	
N	$\rho_0^{\text{ComV@R}}$	$\rho^{g^{\text{V@R}}}$	$\rho_0^{\text{ComV@R}}$	$\rho^{g^{\text{V@R}}}$	$\rho_0^{\text{ComV@R}}$	$\rho^{g^{\text{V@R}}}$
2	27.83	35.71	4.54	7.00	4.86	4.95
4	28.83	62.38	9.52	67.11	4.87	5.09
8	29.53	87.97	12.10	249.49	4.87	5.29
16	29.87	256.06	12.07	$3.97 \cdot 10^3$	4.87	5.57
32	30.20	$1.20 \cdot 10^5$	13.61	$9.15 \cdot 10^5$	4.87	5.88
64	30.32	$4.75 \cdot 10^8$	12.59	$2.16 \cdot 10^9$	4.87	15.06
128	29.88	$5.39 \cdot 10^{13}$	11.01	$3.35 \cdot 10^{13}$	4.87	$4.43 \cdot 10^6$
256	29.83	$1.42 \cdot 10^{22}$	11.17	$1.82 \cdot 10^{23}$	4.87	$4.04 \cdot 10^{13}$
512	29.80	$9.72 \cdot 10^{33}$	10.51	$6.37 \cdot 10^{34}$	4.87	$7.00 \cdot 10^{26}$
	Call		Barrier payoff		Smooth payoff	
N	$\rho_0^{\text{ComAV@R}}$	$\rho^{g^{\text{AV@R}}}$	$\rho_0^{\text{ComAV@R}}$	$\rho^{g^{\text{AV@R}}}$	$\rho_0^{\text{ComAV@R}}$	$\rho^{g^{\text{AV@R}}}$
2	36.05	46.74	13.48	10.07	4.91	5.04
4	36.26	82.96	25.05	120.30	4.92	5.21
8	37.03	253.11	31.71	446.67	4.92	5.45
16	36.83	852.84	34.78	$2.41 \cdot 10^4$	4.92	5.89
32	37.45	$2.88 \cdot 10^6$	37.46	$3.15 \cdot 10^6$	4.92	9.64
64	37.52	$5.46 \cdot 10^{10}$	39.73	$1.96 \cdot 10^{11}$	4.92	$5.33 \cdot 10^3$
128	36.57	$6.87 \cdot 10^{17}$	40.09	$3.49 \cdot 10^{17}$	4.92	$9.48 \cdot 10^{10}$
256	36.88	$2.34 \cdot 10^{27}$	44.44	$4.32 \cdot 10^{28}$	4.92	$7.61 \cdot 10^{19}$
512	37.79	$8.16 \cdot 10^{44}$	43.23	$2.05 \cdot 10^{44}$	4.91	$6.48 \cdot 10^{37}$

7.4.2 Approximation results for composed risk measures and scaled g–expectations

In this section we pick up an idea from Romanovski who proposes to replace the size of the discrete time steps $h = \frac{T}{N}$ by the constant T in $g^{\text{V@R}}$ and $g^{\text{AV@R}}$ from Section 7.2 in order to make the drivers "constant regardless of time scaling" ([120, pages 14, 16]). The resulting driver for value at risk is

$$g^{\text{V@R}}_{\text{Romanovski}}(t,y,z) = \frac{\Phi^{-1}(1-\lambda)}{\sqrt{T}}\,|z|$$

and for average value at risk we get

$$g^{\text{AV@R}}_{\text{Romanovski}}(t,y,z) = \frac{1}{\lambda\sqrt{2\pi T}}\exp\left(-\frac{\Phi^{-1}(1-\lambda)^2}{2}\right)|z|\,.$$

Table 7.4: Composed risk measures and risk measures from g-expectations - Romanovski.

	Call		Barrier payoff		Smooth payoff	
N	$\rho_0^{\text{ComV@R}}$	$\rho^{g^{\text{V@R}}_{\text{Romanovski}}}$	$\rho_0^{\text{ComV@R}}$	$\rho^{g^{\text{V@R}}_{\text{Romanovski}}}$	$\rho_0^{\text{ComV@R}}$	$\rho^{g^{\text{V@R}}_{\text{Romanovski}}}$
2	27.83	24.71	4.54	4.14	4.86	4.86
4	28.83	27.91	9.52	14.44	4.87	4.86
8	29.53	29.03	12.10	21.93	4.87	4.86
16	29.87	29.62	12.07	25.97	4.87	4.86
32	30.20	29.90	13.61	27.01	4.87	4.86
64	30.32	30.17	12.59	27.22	4.87	4.87
128	29.88	30.21	11.01	28.69	4.87	4.87
256	29.83	30.37	11.17	30.57	4.87	4.87
512	29.80	30.94	10.51	32.93	4.87	4.87

	Call		Barrier payoff		Smooth payoff	
N	$\rho_0^{\text{ComAV@R}}$	$\rho^{g^{\text{AV@R}}_{\text{Romanovski}}}$	$\rho_0^{\text{ComAV@R}}$	$\rho^{g^{\text{AV@R}}_{\text{Romanovski}}}$	$\rho_0^{\text{ComAV@R}}$	$\rho^{g^{\text{AV@R}}_{\text{Romanovski}}}$
2	36.05	31.24	13.48	5.81	4.91	4.92
4	36.26	35.62	25.05	22.81	4.92	4.92
8	37.03	36.65	31.71	35.78	4.92	4.92
16	36.83	37.32	34.78	41.84	4.92	4.92
32	37.45	37.64	37.46	41.69	4.92	4.92
64	37.52	38.27	39.73	41.63	4.92	4.92
128	36.57	38.21	40.09	47.34	4.92	4.92
256	36.88	38.80	44.44	53.71	4.92	4.92
512	37.79	41.23	43.23	64.38	4.91	4.92

Figure 7.1: Composed risk measures and risk measures from g-expectations - Romanovski.

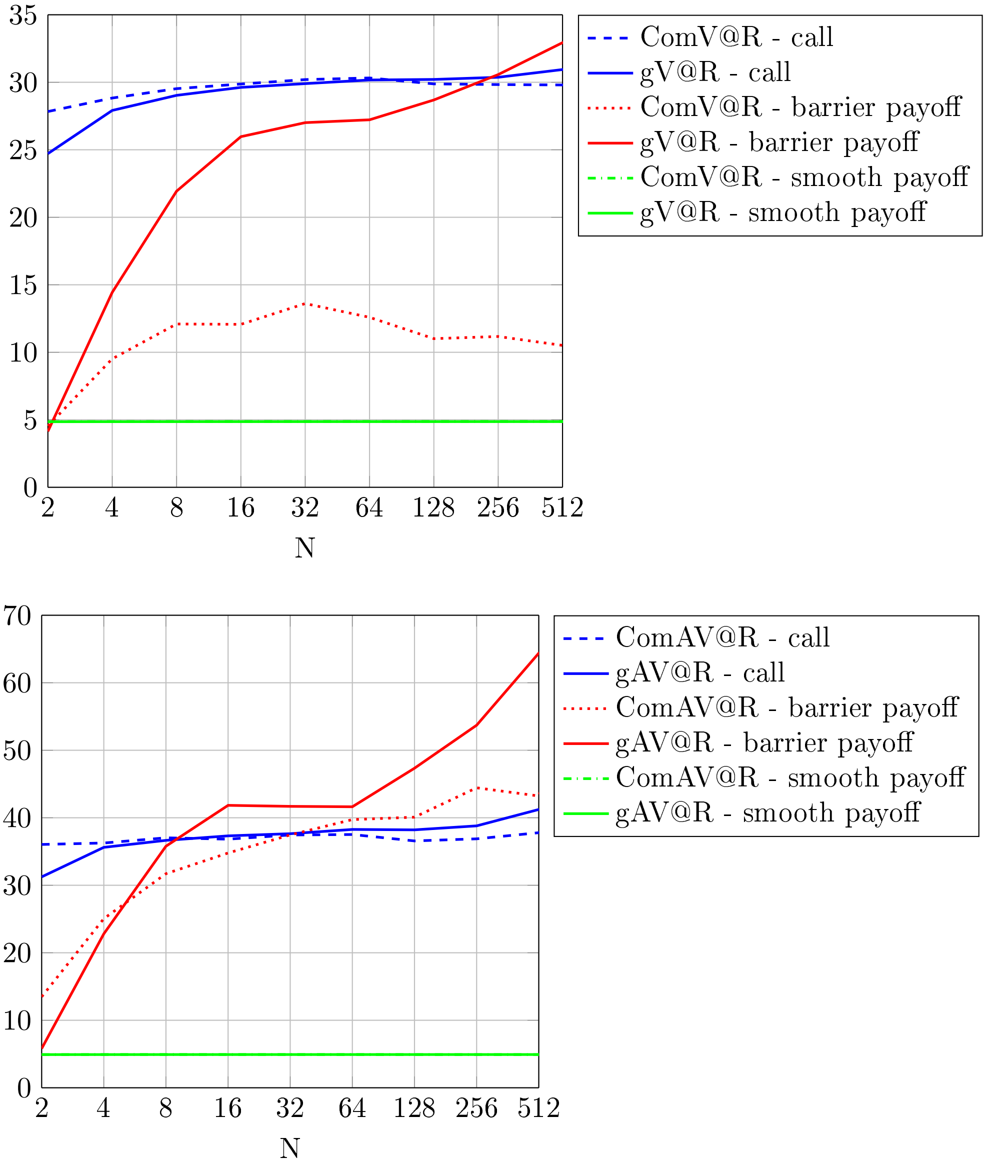


Table 7.4 displays $\rho_0^{\text{ComV@R}}$ ($\rho_0^{\text{ComAV@R}}$) and $\rho^{g^{\text{V@R}}_{\text{Romanovski}}}$ ($\rho^{g^{\text{AV@R}}_{\text{Romanovski}}}$) for $N = 2$ to $N = 512$ time steps. Figure 7.1 graphically displays the results from Table 7.4. We plot g-expectations with solid lines and composed risk measures with dotted lines. The results for the call are shown in blue, the results for the barrier payoff in red and the smooth payoff is displayed in green. Figure 7.1 indicates that the way of subsequently scaling the respective drivers g as proposed by Romanovski [120] only works for payoffs that are "smooth enough"; see the results for call and smooth payoff which confirm the results from [120, Section 3.2]. For the barrier payoff, however, $\rho_0^{\text{ComV@R}}$ and $\rho^{g^{\text{V@R}}_{\text{Romanovski}}}$ do not approach, but rather diverge for greater N. This weakness is also discussed in Figure 7.3

below.

7.4.3 Approximation results for composed scaled risk measures and g–expectations

Let us now consider how the composed scaled dynamic risk measures from Section 7.3 converge to CDRMs solving the corresponding BSDEs. In more specific terms, we calculate composed risk measures in the sense of Stadje [126] and the corresponding solutions to BSDEs with respective drivers, i.e., we compare

$$\rho_0^{\mathrm{ComScaV@R}}(X)$$

to $\rho^{g_{\mathrm{Stadje}}^{\mathrm{V@R}}}(X)$ with

$$g_{\mathrm{Stadje}}^{\mathrm{V@R}}(t,y,z) = \Phi^{-1}(1-\lambda)\,|z|\,,$$

and we compare

$$\rho_0^{\mathrm{ComScaAV@R}}(X)$$

to $\rho^{g_{\mathrm{Stadje}}^{\mathrm{AV@R}}}(X)$ with driver

$$g_{\mathrm{Stadje}}^{\mathrm{AV@R}}(t,y,z) = \frac{1}{\lambda\sqrt{2\pi}}\exp\left(-\frac{\Phi^{-1}(1-\lambda)^2}{2}\right)|z|\,,$$

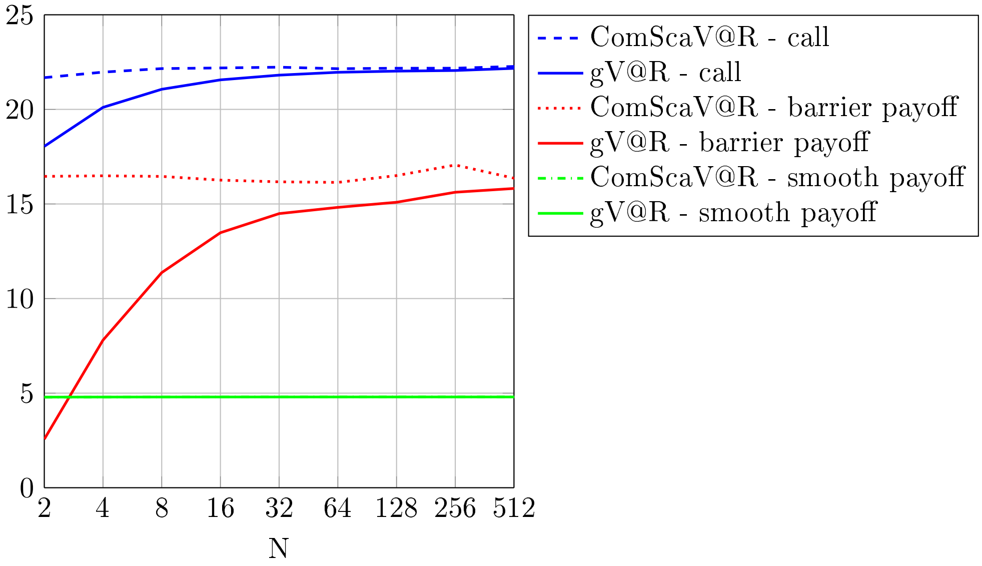
for the smooth payoff, the call and the barrier payoff.

Table 7.5: Composed scaled risk measures and risk measures from g-expectations - Stadje.

	Call		Barrier payoff		Smooth payoff	
N	$\rho_0^{\text{ComScaV@R}}$	$\rho^{g^{\text{V@R}}_{\text{Stadje}}}$	$\rho_0^{\text{ComScaV@R}}$	$\rho^{g^{\text{V@R}}_{\text{Stadje}}}$	$\rho_0^{\text{ComScaV@R}}$	$\rho^{g^{\text{V@R}}_{\text{Stadje}}}$
2	21.68	18.04	16.46	2.57	4.79	4.79
4	21.97	20.11	16.49	7.80	4.79	4.79
8	22.16	21.06	16.46	11.37	4.79	4.80
16	22.19	21.56	16.26	13.48	4.79	4.80
32	22.23	21.81	16.17	14.49	4.80	4.80
64	22.15	21.96	16.14	14.82	4.80	4.80
128	22.17	22.02	16.50	15.09	4.80	4.80
256	22.16	22.07	17.06	15.62	4.80	4.80
512	22.27	22.17	16.36	15.82	4.80	4.80

	Call		Barrier payoff		Smooth payoff	
N	$\rho_0^{\text{ComScaAV@R}}$	$\rho^{g^{\text{AV@R}}_{\text{Stadje}}}$	$\rho_0^{\text{ComScaAV@R}}$	$\rho^{g^{\text{AV@R}}_{\text{Stadje}}}$	$\rho_0^{\text{ComScaAV@R}}$	$\rho^{g^{\text{AV@R}}_{\text{Stadje}}}$
2	27.92	22.03	42.58	3.49	4.82	4.83
4	26.29	24.75	24.39	11.53	4.82	4.84
8	26.63	25.83	22.21	17.24	4.83	4.84
16	26.79	26.39	22.80	20.44	4.83	4.84
32	26.81	26.66	22.76	21.58	4.83	4.84
64	26.63	26.86	24.82	21.87	4.83	4.84
128	26.73	26.92	25.03	22.66	4.83	4.84
256	26.79	27.01	27.49	23.78	4.83	4.84
512	27.23	27.31	25.11	24.86	4.83	4.84

Figure 7.2: Composed scaled risk measures and risk measures from g-expectations - Stadje.



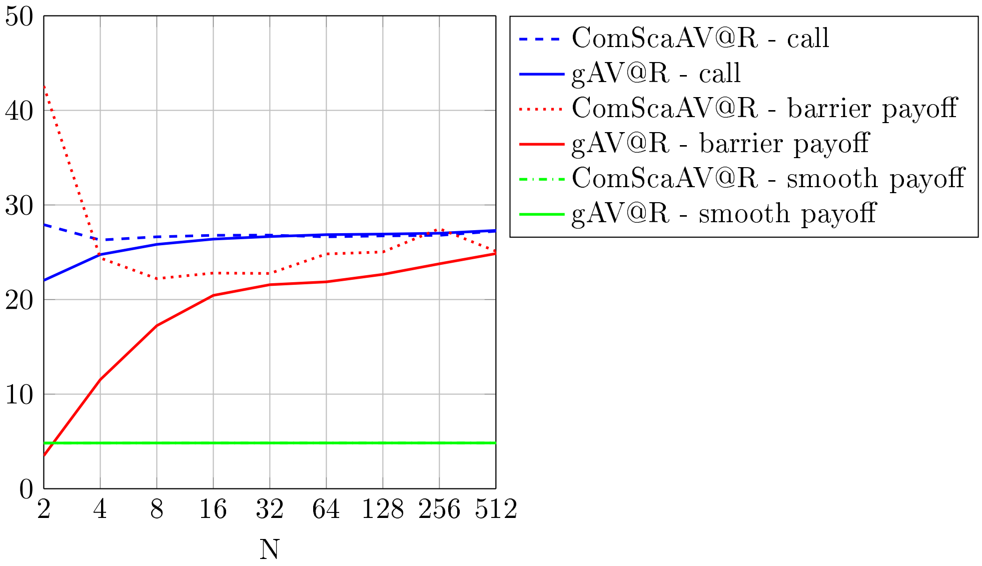
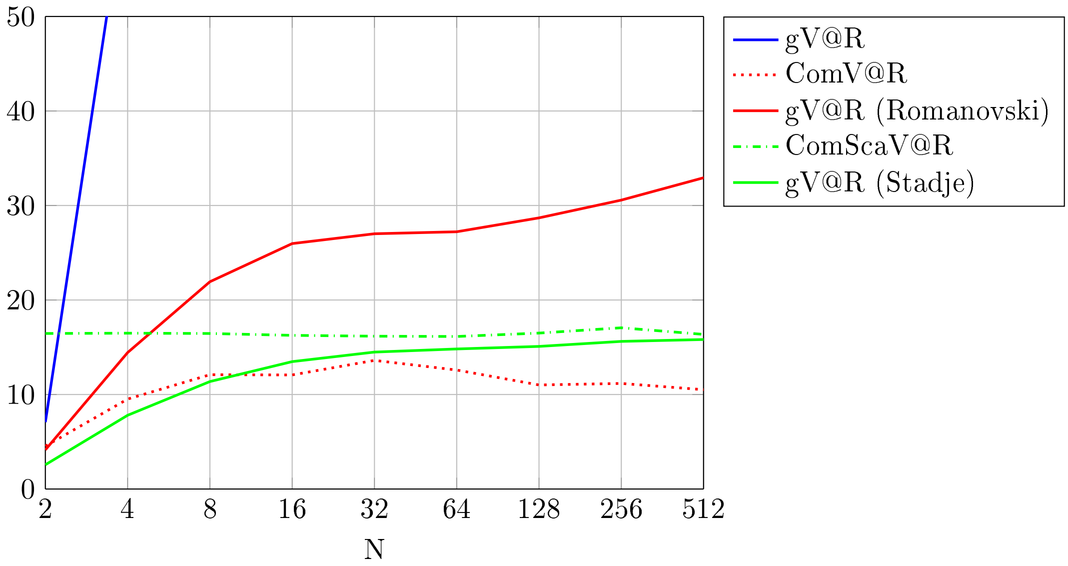


Table 7.5 and Figure 7.2 show $\rho_0^{\text{ComScaV@R}}$ ($\rho_0^{\text{ComScaAV@R}}$) and $\rho^{g^{\text{V@R}}_{\text{Stadje}}}$ ($\rho^{g^{\text{AV@R}}_{\text{Stadje}}}$) for $N = 2$ to $N = 512$ time steps. Again, we plot g-expectations with solid lines and composed scaled risk measures with dotted lines, and the results for the call are shown in blue, the results for the barrier payoff in red and the smooth payoff is displayed in green. We observe a much improved convergence for an increasing number of time steps N from discrete-time composed risk measures to their corresponding continuous-time g-expectations in any tested payoff if we compare Figure 7.2 to Figure 7.1.

In Figure 7.3 we zoom in on the V@R-based risk measures for the barrier payoff.

Figure 7.3: V@R-based risk measures for the barrier payoff.



Firstly, note that the solid blue line displays the g-expectation at $t = 0$ for the naive approach when we do not consider scaling ComV@R or $g^{\text{V@R}}$, respectively. Obviously, the corresponding CDRM $\rho^{g^{\text{AV@R}}}$ "explodes" for bigger N which gives numerical evidence for [126, Proposition 5.1] as V@R is a coherent risk measure for Gaussian random variables. Further, note the clear difference in the convergence behavior for an increasing number of time steps N of Romanovski's approach to the scaling as proposed by Stadje.

7.7 Remark. We have seen the necessity of scaling composed risk measures or the corresponding BSDE-drivers g in order to ensure convergence of ρ_0^{ComRM} to ρ^g for an increasing number of time steps. We, however, strongly recommend to scale composed risk measures as proposed by Stadje [126] as subsequently scaling g as proposed by Romanovski [120] does not imply convergence for non-smooth payoffs; see Figure 7.3. The decisive effect of scaling dynamic risk measures, however, is getting the factor $\sqrt{h}$ in the representation of ComRM to h. With this insight, one might study alternative ways of scaling ComRM, for example, consider

$$\begin{aligned} \rho_{t_N}^{\text{ComSqua}}(X) &:= \rho_{t_N}(X) = -X \\ \rho_{t_k}^{\text{ComSqua}}(X) &:= \tilde{\rho}_{t_k,t_{k+1}}(-\rho_{t_{k+1}}^{\text{ComSqua}}(X)), \quad t_k \le t_{k_{N-1}}, \end{aligned} \tag{ComSquaRM}$$

with

$$\tilde{\rho}_{t_k,t_{k+1}}(X) := \mathbb{E}[-X \mid \mathcal{F}_{t_k}] + h\left(\rho_{t_k,t_{k+1}}\left(X - \mathbb{E}[X \mid \mathcal{F}_{t_k}]\right)\right)^2.$$

For risk measures of the form

$$\rho^{\lambda}(X) := \mathbb{E}[-X] + b(\lambda)\sigma(X),$$

we get

$$\rho_{t_{N-1}}^{\text{ComSqua}}(Y_{t_N}^N) = \mathbb{E}[Y_{t_N}^N \mid \mathcal{F}_{t_{N-1}}] + h\left(\rho_{t_{N-1},t_N}\left(Y_{t_N}^N - \mathbb{E}[Y_{t_N}^N \mid \mathcal{F}_{t_{N-1}}]\right)\right)^2,$$

and the corresponding BSDE-driver is

$$g^b_{\text{Squared}}(t,y,z) = b(\lambda)^2 h z^2.$$

For the same parameter setting as before, composed squared risk measures and the corresponding g-expectations at $t = 0$ are displayed in Table 7.6.

Table 7.6: Composed risk measures and risk measures from g-expectations - squared.

	Call		Barrier payoff		Smooth payoff	
N	$\rho_0^{\text{ComSquaV@R}}$	$\rho^{g^{\text{V@R}}_{\text{Squared}}}$	$\rho_0^{\text{ComSquaV@R}}$	$\rho^{g^{\text{V@R}}_{\text{Squared}}}$	$\rho_0^{\text{ComSquaV@R}}$	$\rho^{g^{\text{V@R}}_{\text{Squared}}}$
2	8.24	6.28	2.48	0.32	4.65	4.64
4	7.79	6.72	2.00	0.87	4.64	4.64
8	7.58	7.01	1.81	1.20	4.64	4.64
16	7.47	7.19	1.71	1.43	4.64	4.64
32	7.42	7.28	1.69	1.56	4.64	4.64
64	7.39	7.32	1.67	1.61	4.64	4.64
128	7.39	7.35	1.65	1.61	4.64	4.64
256	7.38	7.36	1.67	1.67	4.64	4.64
512	7.38	7.37	1.63	1.62	4.64	4.64
	Call		Barrier payoff		Smooth payoff	
N	$\rho_0^{\text{ComSquaAV@R}}$	$\rho^{g^{\text{AV@R}}_{\text{Squared}}}$	$\rho_0^{\text{ComSquaAV@R}}$	$\rho^{g^{\text{AV@R}}_{\text{Squared}}}$	$\rho_0^{\text{ComSquaAV@R}}$	$\rho^{g^{\text{AV@R}}_{\text{Squared}}}$
2	8.84	6.63	3.20	0.35	4.65	4.65
4	8.07	6.91	2.34	0.94	4.64	4.64
8	7.70	7.11	1.95	1.26	4.64	4.64
16	7.53	7.24	1.77	1.47	4.64	4.64
32	7.45	7.30	1.71	1.58	4.64	4.64
64	7.40	7.34	1.68	1.62	4.64	4.64
128	7.40	7.36	1.65	1.61	4.64	4.64
256	7.38	7.37	1.68	1.67	4.64	4.64
512	7.38	7.37	1.63	1.62	4.64	4.64

Note that the absolute values of the displayed risk measures for the call and the barrier payoff in Table 7.6 appear to be too low for practical reasons; see Table 7.2. We, however, emphasize that, from a theoretical point of view, (ComSquaRM) represents a suitable alternative to (ComScaRM) as it provides composed risk measures that converge to the respective robust g-expectations; in particular, the g-expectations do not diverge for an increasing number of time steps N as in Table 7.3.

CHAPTER 8

Conclusion

In the first part of this thesis we considered utility-based risk measures. Since this concept brings together two initially separate fields of research, namely risk measures and preference theory, we provided sufficient theoretical foundations and the basic literature on both topics in Chapters 2 and 3.

The introduction of *optimal expected utility* (OEU) risk measures in Chapter 4 is a main contribution of this thesis. We pointed out that OEU perfectly implements our plan to translate the idea of investors' utility captured by u to risk evaluations of financial positions as it is a convex risk measure for most commonly used utility functions, such as power, logarithmic and exponential utilities and because we were able to relate properties of u to properties of ρ^u. The fact that OEU strictly distinguishes between "cash" and "utility" quantities also facilitates interpretation of OEU. We showed that, to the best of our knowledge, OEU is the only existing utility-based risk measure that is non-trivial and coherent if the utility function features constant relative risk aversion. We also developed a method to recover the utility function from a given OEU and provided a dual representation of optimal expected utility risk measures. Some exemplary applications of OEU demonstrating that (a) it attains reasonable values for any characteristics of Bernoulli-type payoffs, (b) it is sensitive with respect to extreme events, and (c) it also suits for detecting risks associated with heavy tailed distributions concluded this chapter.

Furthermore, we introduced *implied risk aversion* as a rating system for retail structured products based on OEU in Chapter 5. The good suitability of implied risk aversion for classifying RSPs rests above all on the fact that OEU takes into account the entire distribution (we say that it is fully supported) of a financial position; by contrast, V@R and AV@R are not fully supported. Whereas standard V@R-based ratings have some significant weaknesses with respect to classifying RSPs, implied risk aversion suits for a more sophisticated view on the risk and return potential of RSPS and is able to aid retail clients (a) to select the most attractive products from a given universe, and (b) to assess whether a given product is attractive per se. Moreover, implied risk aversion can easily be interpreted in terms of an individual investor's risk aversion. For comparability reasons, we simulated payoff profiles of short and mid term warrants on the German blue-chip index DAX® at their respective maturity date. Due to our theoretical and empirical findings, we suggest to providers of rating information to implement an alternative rating system for RSPs on the basis of implied risk aversion. To our knowledge there is no other publication which proposes to consider the risk parameter of any well-known risk measure to get a neutral (indifferent) risk evaluation. Another application for this approach might be the calculation of the permitted proportions of call and put options for a portfolio consisting of a delta one product and additional options on the same underlying for given RSP-based risk classes. In this way, investors with a clearly defined risk appetite can, for example, define portfolio strategies that are aimed at outperforming or protecting a given underlying.

The third part of this thesis focused on time consistency of dynamic risk measures. In Chapter 6 we set up risk measures on the space of distributions. We showed that OEU is concave on the space of distributions, and pointed out that, in the literature, this is a controversial, but also justifiable characteristic of this type of functionals. Following [137], we presented the theory of acceptance and rejection consistency of dynamic risk measures, and we slightly generalized Weber's findings by considering dynamic risk measures with time-dependent parameters.

Finally, in Chapter 7 we brought together several different studies on convergence behavior of recursively composed dynamic risk measures in discrete time settings to corresponding g-expectations which can be interpreted as dynamic risk measures in continuous time. Our main outcome was that we have to distinguish between two classes of dynamic risk measures with respect to their convergence behavior from discrete to continuous time: While risk measures of the first class, such as the entropic risk measure, are well suited for convergence in the above sense, risk measures of the second class must be additionally scaled in the recursive construction. We proposed an approach for scaling recursively composed risk measures in the sense of [126] and explained why it is preferable to the approach in [120]. We implemented both, unscaled and scaled recursively composed risk measures and derived drivers of BSDEs leading to corresponding continuous-time risk measures for various example cases. Our work on this topic concludes with an illustration of the convergence behavior of V@R- and AV@R-based composed risk measures for different payoffs. Remark 7.7 provides a starting point for further research on recursively composed risk measures.

Bibliography

[1] B. Acciaio and I. Penner. Dynamic risk measures. In *Advanced Mathematical Methods for Finance*, chapter 1, pages 1–34. Springer, 2011.

[2] B. Acciaio and G. Svindland. Are law-invariant risk functions concave on distributions? *Dependence Modeling*, 1:54–64, 2013.

[3] C. Acerbi, C. Nordio, and C. Sirtori. Expected shortfall as a tool for financial risk management. *Working paper*, 2001. Available online at: `http://arxiv.org/pdf/cond-mat/0102304v1`.

[4] C. Acerbi and D. Tasche. On the coherence of expected shortfall. *Journal of Banking and Finance*, 26:1487–1503, 2002.

[5] B. Armbruster and E. Delage. Decision making under uncertainty when preference information is incomplete. *Management Science*, 61(1):111–128, 2015.

[6] K. J. Arrow. *Essays in the Theory of Risk-Bearing*. Markham, Chicago, 3rd edition, 1971.

[7] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Thinking coherently. *Risk Magazine*, 10(11):68–72, 1997.

[8] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.

[9] P. Artzner, F. Delbaen, J.-M. Eber, D. Heath, and H. Ku. Coherent multiperiod risk adjusted values and bellman's principle. *Annals of Operations Research*, 152(1):5–22, 2007.

[10] B. Baeumer and M. M. Meerschaert. Tempered stable Lévy motion and transit superdiffusion. *Journal of Computational and Applied Mathematics*, 233(10):2438–2448, 2010.

[11] O. E. Barndorff-Nielsen and S. Levendorskii. Feller processes of normal inverse Gaussian type. *Quantitative Finance*, 1:318–331, 2001.

[12] O. E. Barndorff-Nielsen and N. Shephard. Normal modified stable processes. Technical report, University of Oxford, 2001.

[13] P. Barrieu and N. El Karoui. Pricing, hedging and optimally designing derivatives via minimization of risk measures. In R. Carmona, editor, *Indifference Pricing: Theory and Applications*. Princeton University Press, 2008.

[14] R. B. Barsky, M. S. Kimball, F. T. Juster, and M. D. Shapiro. Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement survey. *The Quarterly Journal of Economics*, 112(2):537–579, 1997.

[15] R. Baule and P. Blonski. The demand for warrants and issuer pricing strategies. *Journal of Futures Markets, forthcoming*, 2015. Available online at: `http://onlinelibrary.wiley.com/doi/10.1002/fut.21703/pdf`.

[16] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York, 2011.

[17] F. Bellini and V. Bignozzi. On elicitable risk measures. *Quantitative Finance*, 15(5):725–733, 2015.

[18] A. Ben-Tal and M. Teboulle. Expected utility, penalty functions and duality in stochastic nonlinear programming. *Management Science*, 32:1445–1466, 1986.

[19] A. Ben-Tal and M. Teboulle. An old-new concept of convex risk measures: The optimized certainty equivalent. *Mathematical Finance*, 17(3):449–476, 2007.

[20] H. P. Binswanger. Attitudes toward risk: Theoretical implications of an experiment in rural India. *The Economic Journal*, pages 867–890, 1981.

[21] F. Black and M. Scholes. The pricing of options and corporate liabilities. *The Journal of Political Economy*, 87:637–659, 1973.

[22] T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327, 1986.

[23] B. Bouchard and N. Touzi. Discrete time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Processes and Their Applications*, 111:175–206, 2004.

[24] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel. *Time Series Analysis: Forecasting and Control*. Prentice-Hall, Englewood Cliffs, 1994.

[25] W. Breuer and A. Perst. Retail banking and behavioral financial engineering: The case of structured products. *Journal of Banking and Finance*, 31(3):827–844, 2007.

[26] P. Briand, B. Delyon, and J. Mémin. Donsker-type theorem for BSDEs. *Electronic Communications in Probability*, 6:1–14, 2001.

[27] J. Cao and M. O. Rieger. Risk classes for structured products: Mathematical aspects and their implications on behavioral investors. *Annals of Finance*, 9(2):167–183, 2013.

[28] P. Carr, H. Geman, D. R. Madan, and M. Yor. The fine structure of asset returns: An empirical investigation. *Journal of Business*, 75:305–332, 2002.

[29] C. Celerier and B. Vallee. The motives for financial complexity: An empirical investigation. *HEC Paris Research Paper No. FIN-2013-1013*, 2014. Available online at: `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2289890`.

[30] A. H. Chen and J. W. Kensinger. An analysis of market-index certificates of deposit. *Journal of Financial Services Research*, 4:93–110, 1990.

[31] K. C. Chen and R. S. Sears. Pricing the SPIN. *Financial Management*, 19:36–47, 1990.

[32] P. Cheridito, F. Delbaen, and M. Kupper. Coherent and convex monetary risk measures for bounded càdlàg processes. *Stochastic Processes and Their Applications*, 112(1):1–22, 2004.

[33] P. Cheridito, F. Delbaen, and M. Kupper. Coherent and convex monetary risk measures for unbounded càdlàg processes. *Finance and Stochastics*, 9(3):369–387, 2005.

[34] P. Cheridito, F. Delbaen, and M. Kupper. Dynamic monetary risk measures for bounded discrete-time processes. *Electronic Journal of Probability*, 11(3):57–106, 2006.

[35] P. Cheridito and M. Kupper. Composition of time-consistent dynamic monetary risk measures in discrete time. *International Journal of Theoretical and Applied Finance*, 14(1):137–162, 2011.

[36] P. Cheridito and M. Stadje. Time-inconsistency of VaR and time-consistent alternatives. *Finance Research Letters*, 6:40–46, 2009.

[37] P.-A. Chiappori and M. Paiella. Relative risk aversion is constant: Evidence from panel data. *Journal of the European Economic Association*, 9(6):1021–1052, 2011.

[38] S. N. Cohen and R. J. Elliott. Backward stochastic differential equations and nearly-time-consistent nonlinear expectations. *SIAM Journal of Control and Optimization*, 49:125–139, 2011.

[39] R. Cont. Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1:223–236, 2001.

[40] J. Daníelsson. *Financial Risk Forecasting: The Theory and Practice of Forecasting Market Risk With Implementation in R and Matlab*, volume 588. John Wiley & Sons, 2011.

[41] J. Daníelsson, B. N. Jorgensen, G. Samorodnitsky, M. Sarma, and C. G. de Vries. Subadditivity re-examined: The case for value-at-risk . *Working paper, London School of Economics*, 2005.

[42] B. de Finetti. *Sul Concetto di Media*. Istituto italiano degli attuari, 1931.

[43] F. Delbaen. Coherent risk measures. *Cattedra Galileiana, Scuola Normale Superiore, Classe di Scienze, Pisa*, 2000.

[44] F. Delbaen. Coherent risk measures on general probability spaces. In *Advances in Finance and Stochastics*, volume 6, pages 429–447. Springer, 2002.

[45] F. Delbaen. Risk measures or measures that describe risk? Available online at: `http://www.math.ethz.ch/~delbaen/ftp/preprints/Risk-executives.ps`, 2008.

[46] K. Detlefsen and G. Scandolo. Conditional and dynamic convex risk measures. *Finance and Stochastics*, 9(4):539–561, 2005.

[47] Deutscher Derivate Verband. Welche Risikopräferenz haben Sie beim Kauf eines Zertifikats? Available online at: `http://www.derivateverband.de/DE/MediaLibrary/Document/PM/15%2003%2010%20PM%20DDV-Trend%20des%20Monats%20M%C3%A4rz%202015.pdf`, 2015.

[48] J. Dhaene, M. J. Goovaerts, and R. Kaas. Economic capital allocation derived from risk measures. *North American Actuarial Journal*, 7(2):44–56, 2003.

[49] J. Dhaene, R. J. A. Laeven, S. Vanduffel, G. Darkiewicz, and M. J. Goovaerts. Can a coherent risk measure be too subadditive? *Journal of Risk and Insurance*, 75(2):365–386, 2008.

[50] B. Döhrer, L. Johanning, N. Steiner, and A. Völkle. Risikoberechnungs- und Klassifizierungsmethoden nach KIID und EDG. Available online at: `http://www.derivateverband.de/ENG/Publications/StudiesAndReports`, 2012.

[51] B. Döhrer, L. Johanning, N. Steiner, and A. Völkle. Issuer margins for structured products in Germany. Available online at: `http://www.derivateverband.de/ENG/Publications/StudiesAndReports`, 2013.

[52] J. Douglas, J. Ma, and P. Protter. Numerical methods for forward-backward stochastic differential equations. *The Annals of Applied Probability*, 6(3):940–968, 1996.

[53] K. Dowd and D. Blake. After VaR: The theory, estimation, and insurance applications of quantile-based risk measures. *Journal of Risk and Insurance*, 73(2):193–229, 2006.

[54] S. Drapeau and M. Kupper. Risk preferences and their robust representation. *Mathematics of Operations Research*, 38(1):28–62, 2013.

[55] D. Duffie and J. Pan. An Overview of Value at Risk. *Journal of Derivatives*, 4:7–49, 1997.

[56] P. Embrechts, A. McNeil, and D. Straumann. Correlation and dependence in risk management: Properties and pitfalls. In M. A. H. Dempster, editor, *Risk Management: Value at Risk and Beyond*, pages 176–223. Cambridge University Press, Cambridge, 2002.

[57] R. F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50:987–1007, 1982.

[58] O. Entrop, A. Schober, and M. Wilkens. The pricing policy of banks for exchange-traded leverage certificates: Interday and intraday effects. Available online at: `http://ssrn.com/abstract=1757710`, 2013.

[59] European Union. DIRECTIVE 2009/65/EC. Available online at: `http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32009L0065`, 2009.

[60] European Union. Regulation (EU) No 575/2013. Available online at: `http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32013R0575`, 2013.

[61] H. Fink, S. Geissel, J. Sass, and F. T. Seifried. Implied risk aversion: An alternative rating system for retail structured products. Preprint, Available online at: `http://ssrn.com/abstract=2651135`, 2015.

[62] H. Fink and S. Mittnik. Pricing Nikkei 225 quanto options via normal tempered stable processes. Working paper, Ludwig Maximilian University of Munich, 2015.

[63] P. C. Fishburn. *Utility Theory for Decision Making.* Wiley, New York, 1970.

[64] P. C. Fishburn. *Nonlinear Preference and Utility Theory.* Johns Hopkins University Press, 1988.

[65] H. Föllmer and T. Knispel. Entropic risk measures: Coherence vs. convexity, model ambiguity, and robust large deviations. *Stochastics and Dynamics*, 11(2-3):333–351, 2011.

[66] H. Föllmer and I. Penner. Convex risk measures and the dynamics of their penalty functions. *Statistics and Decisions*, 24(1):61–96, 2006.

[67] H. Föllmer and A. Schied. Convex Measures of Risk and Trading Constraints. *Finance and Stochastics*, 6:429–447, 2002.

[68] H. Föllmer and A. Schied. Convex and coherent risk measures. In R. Cont, editor, *Encyclopedia of Quantitative Finance*, pages 355–363. John Wiley & Sons, 2010.

[69] H. Föllmer and A. Schied. *Stochastic Finance.* Walter de Gruyter, Berlin New York, 3rd edition, 2011.

[70] H. Föllmer and A. Schied. Probabilistic aspects of finance. *Bernoulli*, 19(4):1306–1326, 2013.

[71] H. Föllmer, A. Schied, and S. Weber. Robust Preferences and Robust Portfolio Choice. *Handbook of Numerical Analysis, XV, Bensoussan & Zhang (Editors), Mathematical Modeling and Numerical Methods in Finance*, pages 29–87, 2009.

[72] H. Föllmer and S. Weber. The axiomatic approach to risk measures for capital determination. *Annual Review of Financial Economics*, 7, 2015.

[73] I. Friend and M. E. Blume. The demand for risky assets. *The American Economic Review*, pages 900–922, 1975.

[74] M. Fritelli and E. Rosazza Gianin. Putting order in risk measures. *Journal of Banking and Finance*, 26:1473–1486, 2002.

[75] M. Frittelli, M. Maggis, and I. Peri. Risk measures on P(R) and Value at Risk with Probability/Loss function. *Mathematical Finance*, 24(3):442–463, 2014.

[76] M. Frittelli and G. Scandolo. Risk measures and capital requirements for processes. *Mathematical finance*, 16(4):589–612, 2006.

[77] S. Geissel, J. Sass, and F. T. Seifried. Optimal expected utility risk measures. Preprint, Available online at: `http://ssrn.com/abstract=2651132`, 2015.

[78] K. Giesecke, T. Schmidt, and S. Weber. Measuring the risk of large losses. *Journal of Investment Management*, 6(4):1–15, 2008.

[79] I. Gilboa and D. Schmeidler. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18:141–153, 1989.

[80] E. Gobet, J.-P. Lemor, and X. Warin. A regression based Monte Carlo method to solve backward stochastic differential equations. *The Annals of Applied Probability*, 15(3):2172–2202, 2005.

[81] C. Gollier. *The Economics of Risk and Time.* MA: MIT Press, Cambridge, 2001.

[82] B. E. Hansen. Autoregressive conditional density estimation. *International Economic Review*, 35:705–730, 1994.

[83] D. Heath. *Back to the future. Plenary lecture.* First World Congress of the Bachelier Finance Society, Paris, 2000.

[84] D. Heath and H. Ku. Pareto equilibria with coherent measures of risk. *Mathematical Finance*, 14(2):163–172, 2004.

[85] B. J. Henderson and N. D. Pearson. The dark side of financial innovation: A case study of the pricing of a retail financial product. *Journal of Financial Economics*, 100:227–247, 2011.

[86] T. Hens and M. O. Rieger. Can utility optimization explain the demand for structured investment products? *Quantitative Finance*, 14(4):673–681, 2014.

[87] K. Janecek. What is a realistic aversion to risk for real-world individual investors? Available online at: `http://www.sba21.com/personal/RiskAversion.pdf`, 2004.

[88] E. Jouini, W. Schachermayer, and N. Touzi. Law invariant risk measures have the fatou property. In *Advances in Mathematical Economics*, pages 49–71. Springer, 2006.

[89] R. Kawai and H. Masuda. Infinite variation tempered stable Ornstein-Uhlenbeck processes with discrete observations. *Communications in Statistics - Simulation and Computation*, 41(1):125–139, 2012.

[90] Y. S. Kim, R. Giacometti, S. T. Rachev, F. J. Fabozzi, and D. Mignacca. Measuring financial risk and portfolio optimization with a non-Gaussian multivariate model. *Annals of Operations Research*, 201(1):325–343, 2012.

[91] Y. S. Kim and D. Volkmann. NTS copula and finance. *Applied Mathematics Letters*, 26:676–680, 2013.

[92] M. S. Kimball. Precautionary saving in the small and in the large. *Econometrica*, pages 53–73, 1990.

[93] M. S. Kimball. Standard risk aversion. *Econometrica*, pages 589–611, 1993.

[94] M. S. Kimball, C. R. Sahm, and M. D. Shapiro. Imputing risk tolerance from survey responses. *Journal of the American Statistical Association*, 103(483):1028–1038, 2008.

[95] S. G. Kou, X. H. Peng, and C. C. Heyde. External risk measures and basel accords. *Mathematics of Operations Research*, 38(3):393–417, 2013.

[96] D. M. Kreps and E. L. Porteus. Temporal resolution of uncertainty and dynamic choice theory. *Econometrica*, 46:185–200, 1978.

[97] P. A. Krokhmal. Higher moment coherent risk measures. *Quantitative Finance*, 7(4):373–387, 2007.

[98] M. Kupper and W. Schachermayer. Representation results for law invariant time consistent functions. *Mathematics and Financial Economics*, 2(3):189–210, 2009.

[99] J. Ma, P. Protter, and J. Yong. Solving forward-backward stochastic differential equations explicitly - a four step scheme. *Probability Theory and Related Fields*, 98(3):339–359, 1994.

[100] F. Maccheroni, M. Marinacci, and A. Rustichini. Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, pages 1447–1498, 2006.

[101] D. B. Madan, P. Carr, and E. C. Chang. The Variance Gamma process and option pricing. *European Finance Review*, 2(1):79–105, 1998.

[102] D. B. Madan and F. Milne. Option pricing with VG martingale components. *Mathematical Finance*, 1(4):39–55, 1991.

[103] D. B. Madan and E. Seneta. The VG model for share market returns. *Journal of Business*, 63:511–524, 1990.

[104] H. Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952.

[105] A. Mas-Colell, M. D. Whinston, and J. R. Green. *Microeconomic Theory.* Oxford University Press, Oxford, 1995.

[106] R. C. Merton. Rational theory of option pricing. *Bell Journal of Economics and Management Science*, 4:141–83, 1973.

[107] D. J. Meyer and J. Meyer. Relative risk aversion: What do we know? *The Journal of Risk and Uncertainty*, 31(3):243–262, 2005.

[108] A. Müller. Certainty equivalents as risk measures. *Brazilian Journal of Probability and Statistics*, 21:1–12, 2007.

[109] M. Nagumo. Über eine Klasse der Mittelwerte. *Japanese Journal of Mathematics*, 7:71–79, 1930.

[110] D. Paravisini, V. Rappoport, and E. Ravina. Risk aversion and wealth: Evidence from person-to-person lending portfolios. Technical report, National Bureau of Economic Research, 2010.

[111] S. Peng. Backward SDE and related g-expectations. In N. El Karoui and S. Mazliak, editors, *Backward Stochastic Differential Equations. Pitman Research Notes in Mathematics Series*, volume 364, pages 141–159. Addison Wesley Longman Limited, 1997.

[112] I. Penner. Dynamic Convex Risk Measures: Time Consistency, Prudence, and Sustainability. *Ph.D. thesis, Humboldt-Universität Berlin*, 2007.

[113] R. A. Pollak. Consistent Planning. *The Review of Economic Studies*, 35(2):201–208, 1968.

[114] J. W. Pratt. Risk aversion in the small and in the large. *Econometrica*, 32(1-2):122–136, 1964.

[115] S. T. Rachev. *Handbook of Heavy Tailed Distributions in Finance.* Elsevier, Amsterdam, Boston, 2003.

[116] F. Riedel. Dynamic coherent risk measures. *Stochastic Processes and their Applications*, 112(2):185–200, 2004.

[117] R. T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–41, 2000.

[118] R. T. Rockafellar and S. Uryasev. Conditional value-at-risk for general loss distributions. *Journal of Banking and Finance*, 26:1443–1471, 2002.

[119] R. T. Rockafellar, S. Uryasev, and M. Zabarankin. Generalized deviations in risk analysis. *Finance and Stochastics*, 10(1):51–74, 2006.

[120] I. Romanovski. Time consistency in risk measures. *Master Thesis, St Cross College, University of Oxford*, 2012.

[121] E. Rosazza Gianin. Risk Measures via g-expectations. *Insurance Mathematics and Economics*, 39:19–34, 2006.

[122] S. Rosetto and J. van Bommel. Endless leverage certificates. *Journal of Banking and Finance*, 33(8):1543–1553, 2009.

[123] L. J. Savage. *The Foundations of Statistics*. Wiley, New York, 1954.

[124] P. Schmitz and M. Weber. Buying and selling behaviour of individual investors in option-like securities. *Die Betriebswirtschaft*, 72:409–426, 2012.

[125] S. Schroff, S. Meyer, and H.-P. Burghof. Retail investor information demand - speculating and investing in structured products. *The European Journal of Finance*, 2015. Available online at: `http://www.tandfonline.com/doi/abs/10.1080/1351847X.2015.1020392#.Vctejfk1-Uk`.

[126] M. Stadje. Extending dynamic convex risk measures from discrete time to continuous time: A convergence approach. *Insurance: Mathematics and Economics*, 47:391–404, 2010.

[127] P. A. Stoimenov and S. Wilkens. Are structured products 'fairly' priced? An analysis of the German market for equity-linked instruments. *Journal of Banking and Finance*, 29:2971–2993, 2005.

[128] R. H. Strotz. Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, 23(3):165–180, 1956.

[129] G. G. Szpiro. Measuring risk aversion: An alternative approach. *The Review of Economics and Statistics*, pages 156–159, 1986.

[130] J. ter Horst and C. Veld. An empirical analysis of the pricing of bank issued options versus options exchange options. *European Financial Management Association*, 14(2):288–314, 2008.

[131] S. Tutsch. Konsistente und kosequente dynamische Risikomaße und das Problem der Aktualisierung. *Ph.D. thesis, Humboldt-Universität Berlin*, 2006.

[132] S. Tutsch. Update rules for convex risk measures. *Quantitative Finance*, 8:833–843, 2008.

[133] A. Vinel and P. A. Krokhmal. Certainty equivalent measures of risk. *Annals of Operations Research*, 2015.

[134] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, 2nd edition, 1947.

[135] M. Wallmeier. Beyond payoff diagrams: How to present risk and return characteristics of structured products. *Financial Markets and Portfolio Management*, 25(3):313–338, 2011.

[136] S. Weber. Measures and Models of Financial Risk. *Ph.D. thesis, Humboldt University Berlin*, 2004.

[137] S. Weber. Distribution-invariant risk measures, information, and dynamic consistency. *Mathematical Finance*, 16(2):419–441, 2006.

[138] S. Weber and T. Schmidt. Alternativen zu Value at Risk. *Zeitschrift für die gesamte Versicherungswissenschaft*, 4:669–690, 2005.

[139] P. Whittle. *Hypothesis Testing in Time Series Analysis.* Almquist and Wicksell, Uppsala, 1951.

[140] S. Wilkens, C. Erner, and K. Röder. The pricing of structured products in Germany. *Journal of Derivatives*, 11:55–69, 2003.

[141] S. Wilkens and A. Stoimenov. The pricing of leverage products: An empirical investigation of the German market for long and short stock index certificates. *Journal of Banking and Finance*, 31:735–750, 2007.

Appendix

A.1 Proofs

A.1 Remark (Under (PH), (C) is equivalent to (S)). We assume that a risk measure ρ is positively homogeneous.

1. Convexity implies that:

$$\rho(X+Y) = \rho\left(\frac{1}{2}2X + \frac{1}{2}2Y\right) \leq \frac{1}{2}\rho(2X) + \frac{1}{2}\rho(2Y) = \rho(X) + \rho(Y).$$

2. Subadditivity implies that for any $0 \leq \lambda \leq 1$:

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

Thus we conclude that, under the assumption of positive homogeneity, (C) is equivalent to (S).

A.2 Remark ($V@R$ is positively homogeneous). For $\gamma > 0$:

$$\begin{aligned} V@R^{\lambda}(\gamma X) &= \inf\{m \in \mathbb{R} : P(m + \gamma X < 0) \leq \lambda\} \\ &= \gamma \inf\left\{\frac{m}{\gamma} \in \mathbb{R} : P\left(\frac{m}{\gamma} + X < 0\right) \leq \lambda\right\} \\ &= \gamma V@R^{\lambda}(X) \end{aligned}$$

A.3 Remark ($V@R$ is not subadditive). Consider the example of an investor selling two options on the same underlying S with the same maturity, namely:

- a call with strike K_{call} at p_{call} and
- a put with strike K_{put} at p_{put},

such that

$$P(S < K_{\text{put}} - p_{\text{put}}) = P(S > K_{\text{call}} + p_{\text{call}}) = 0.0075$$

and

$$P(S < K_{\text{put}} - (p_{\text{call}} + p_{\text{put}})) = P(S > K_{\text{call}} + (p_{\text{call}} + p_{\text{put}})) = 0.006.$$

Set $X_1 := -((K_{\text{put}} - S)^+ - p_{\text{put}})$, $X_2 := -((S - K_{\text{call}})^+ - p_{\text{call}})$. Then $V@R^{0.01}(X_1) \le 0$, $V@R^{0.01}(X_2) \le 0$, but $V@R^{0.01}(X_1 + X_2) > 0$

A.4 Remark ($(\rho^{\text{ent}}_{t,\gamma})_{t\in\{0,\dots,T\}}$ is time consistent).

1. $(\rho^{\text{ent}}_{t,\gamma})_{t\in\{0,\dots,T\}}$ is recursive:

$$\begin{aligned}\rho^{\text{ent}}_{t,\gamma}(-\rho^{\text{ent}}_{t+1,\gamma}(X)) &= \frac{1}{\gamma}\mathbb{E}\left[\exp\left(-\gamma\left(-\frac{1}{\gamma}\ln\mathbb{E}[\exp(-\gamma X) \mid \mathcal{F}_{t+1}]\right)\right)\right] \\ &= \frac{1}{\gamma}\ln\mathbb{E}\left[\mathbb{E}[\exp(-\gamma X) \mid \mathcal{F}_{t+1}] \mid \mathcal{F}_t\right] \\ &= \rho^{\text{ent}}_{t,\gamma}(X)\end{aligned}$$

2. Recursiveness implies time consistency: Since ρ_t is monotone,

$$\rho_t(-\rho_{t+1}(X)) \le \rho_t(-\rho_{t+1}(Y)) \quad \text{if} \quad \rho_{t+1}(X) \le \rho_{t+1}(Y)$$

which, due to recursiveness, leads to

$$\rho_t(X) \le \rho_t(Y).$$

It is easy to proof that time consistency is even equivalent to recursiveness.

A.5 Remark ($V@R$ for normally distributed X). If X is (continuously) normally distributed with variance $\sigma^2(X)$, V@R equals the quantile function of $-X$ for $1-\lambda$:

$$V@R^{\lambda}(X) = \mathbb{E}[-X] + \sigma(X)\Phi^{-1}(1-\lambda).$$

A.6 Remark ($V@R$ is subadditive for normally distributed positions). Let $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ and $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ with $corr(X,Y) = \rho \in [0,1]$. Then

$$\begin{aligned}V@R^{\lambda}(X+Y) &= \mu_X + \mu_Y + \sqrt{\sigma_X^2 + \sigma_Y^2 + 2\rho\sigma_X\sigma_Y}\,\Phi^{-1}(1-\lambda) \\ &\le \mu_X + \mu_Y + \sqrt{\sigma_X^2 + \sigma_Y^2 + 2\sigma_X\sigma_Y}\,\Phi^{-1}(1-\lambda) \\ &= \mu_X + \mu_Y + \sqrt{(\sigma_X + \sigma_Y)^2}\,\Phi^{-1}(1-\lambda) \\ &= V@R^{\lambda}(X) + V@R^{\lambda}(Y)\end{aligned}$$

A.7 Remark (Increments of a geometric Brownian motion are nearly Gaussian). Let $(S_t)_{t\in[0,T]}$ be a geometric Brownian motion

$$dS_t = \mu S_t dt + \sigma S_t dW_t,$$

i.e., for all t, S_t is log-normally distributed with

$$\mathbb{E}[S_t] = s\exp(\mu t) =: m, \quad Var[S_t] = s^2\exp(2\mu t)\left(\exp(\sigma^2 t) - 1\right) =: v.$$

Then, $\ln(S_t)$ is normally distributed with parameters

$$\hat{\sigma}_t^2 = \ln\left(1 + \frac{v}{m^2}\right) = \sigma^2 t$$

and

$$\begin{aligned}\hat{\mu}_t &= \ln\left(\frac{m^2}{\sqrt{v+m^2}}\right),\\ &= \ln\left(\frac{s^2\exp(2\mu t)}{\sqrt{s^2\exp(2\mu t)\exp(\sigma^2 t)}}\right)\\ &= \ln\left(S_0\exp(t(\mu-\frac{1}{2}\sigma^2))\right)\\ &= \ln(S_0) + t(\mu - \frac{1}{2}\sigma^2)\end{aligned}$$

and with Itô's formula

$$d\ln(S_t) = \tilde{\mu}dt + \sigma dW_t, \quad \tilde{\mu} := \mu - \frac{1}{2}\sigma^2.$$

Thus

$$\ln(S_{t+h}) - \ln(S_t) \sim \mathcal{N}\left(\hat{\mu}_{t+h} - \hat{\mu}_t, \hat{\sigma^2_{t+h}} - \hat{\sigma^2_t}\right) = \mathcal{N}\left(\tilde{\mu}h, \sigma^2 h\right),$$

since $\ln(S_{t+h})$, $\ln(S_t)$ are increments of a standard Brownian motion and therefore stationary independent. It follows that

$$S_{t+h} - S_t = S_t\left(\frac{S_{t+h}}{S_t} - 1\right) \overset{h \text{ very small}}{\approx} S_t \ln\left(\frac{S_{t+h}}{S_t}\right) \sim \mathcal{N}\left(S_t\tilde{\mu}h, S_t^2\sigma^2 h\right).$$

A.8 Remark ($AV@R^\lambda(X) = \mu + \sigma\frac{1}{\lambda\sqrt{2\pi}}\exp\left(-\frac{\Phi^{-1}(1-\lambda)^2}{2}\right)$ for $L = -X \sim \mathcal{N}(\mu,\sigma^2)$). Let $-X \sim \mathcal{N}(\mu,\sigma^2)$, $Z \sim \mathcal{N}(0,1)$. From Remark 2.5(c) we know the representation of $AV@R^\lambda(X)$ for continuously distributed X. Then:

$$\begin{aligned}AV@R^\lambda(X) &= \mathbb{E}[-X : -X \geq V@R^\lambda(X)]\\ &= \mathbb{E}[(\mu+\sigma Z) : (\mu+\sigma Z) \geq \mu + \sigma\Phi^{-1}(1-\lambda)]\\ &= \mu + \sigma\mathbb{E}[Z : Z \geq \Phi^{-1}(1-\lambda)]\\ &= \mu + \sigma\frac{1}{\alpha\sqrt{2\pi}}\exp\left(-\frac{\Phi^{-1}(1-\lambda)^2}{2}\right)\end{aligned}$$

A.2 Data

In this section we present additional data to the RSP ratings from Chapter 5.

Table A.1: Products with largest implied risk aversion in the Black-Scholes model.

WKN	Issuer	Strike	Ask	Underlying	Maturity	V@R	γ_0
UZ68WB	UBS	11950	0.41	11828.86	5	10000	15.4231
UZ7M77	UBS	11900	0.59	11830.62	5	10000	15.0835
UZ62LW	UBS	12000	0.28	11829.15	5	10000	14.6950
HY7UPK	Hypo Vereinsbank	11900	0.53	11831.82	4	10000	14.3529
UZ7LV7	UBS	12050	0.18	11828.89	5	10000	13.6851
UZ698B	UBS	11850	0.81	11829.76	5	10000	13.2928
HY7243	Hypo Vereinsbank	12000	0.24	11828.16	4	10000	12.5903
UZ6U6M	UBS	12100	0.11	11829.38	5	10000	12.3819
XM2HAK	Deutsche Bank	12000	0.32	11831.29	5	10000	12.2658
GL15JW	Goldman Sachs	12000	0.327	11831.26	5	10000	11.7561
XM2HAJ	Deutsche Bank	11950	0.47	11830.66	5	10000	11.5091
XM2HAH	Deutsche Bank	11900	0.65	11831.82	5	10000	11.0532
XM2HAL	Deutsche Bank	12050	0.22	11829.58	5	10000	10.9613
HY7244	Hypo Vereinsbank	12050	0.16	11828.96	4	10000	10.7327
HY7UPL	Hypo Vereinsbank	11950	0.4	11828.59	4	10000	10.4227
UZ68WE	UBS	12150	0.068	11828.16	5	10000	10.2577
UZ7CNC	UBS	11800	1.09	11828.09	5	10000	10.2328
PS2SLG	BNP Paribas	12100	0.24	11832.04	7	10000	10.2182
GL15JY	Goldman Sachs	12100	0.145	11831.93	5	10000	10.1782
XM2HAM	Deutsche Bank	12100	0.15	11830.42	5	10000	9.6278
GL15JZ	Goldman Sachs	12100	0.247	11828.59	7	10000	9.3890
CC67N7	Citi	12100	0.25	11829.59	7	10000	9.3238
GL15JU	Goldman Sachs	11900	0.67	11829.52	5	10000	9.1773
GL15JX	Goldman Sachs	12000	0.47	11827.23	7	10000	9.1561
UZ670C	UBS	12200	0.03	11827.14	5	10000	9.0899
PS2SLF	BNP Paribas	12000	0.48	11829.9	7	10000	9.0124
VZ91QW	Vontobel	12125	0.21	11828.96	7	10000	9.0041
VZ91QV	Vontobel	12075	0.31	11831.96	7	10000	8.9334
VZ9CZ2	Vontobel	12050	0.36	11830.44	7	10000	8.9283
CN263Q	Commerzbank	12075	0.31	11831.82	7	10000	8.9151
VZ8ZUT	Vontobel	12100	0.26	11830.14	7	10000	8.7757
PS2SLH	BNP Paribas	12200	0.11	11830.58	7	10000	8.7722
VZ9CZ3	Vontobel	12150	0.18	11832.04	7	10000	8.7207
HY7245	Hypo Vereinsbank	12100	0.1	11825.44	4	10000	8.6373
XM2HAG	Deutsche Bank	11850	0.88	11827.33	5	10000	8.4603
CC8Z2M	Citi	12150	0.18	11828.98	7	10000	8.4280
VZ91QX	Vontobel	12175	0.15	11831.75	7	10000	8.3502
VZ8ZUS	Vontobel	12000	0.5	11832.1	7	10000	8.2765
CN0R7K	Commerzbank	12100	0.27	11831.13	7	10000	8.2757
CN263P	Commerzbank	12025	0.43	11830.44	7	10000	8.2651
CC67N8	Citi	12200	0.12	11832.04	7	10000	8.2642

CN0R7J	Commerzbank	12050	0.37	11829.59	7	10000	8.1950
AP6MBP	Interactive Brokers	12100	0.27	11829.59	7	10000	8.0900
VZ8ZUR	Vontobel	12200	0.12	11829.42	7	10000	8.0597
HY7UPJ	Hypo Vereinsbank	11850	0.79	11825.86	4	10000	7.9905
CN0R7L	Commerzbank	12150	0.19	11829.9	7	10000	7.8940
CN0R7H	Commerzbank	12000	0.5	11829.26	7	10000	7.8620
CN263R	Commerzbank	12125	0.23	11829.28	7	10000	7.8004
VZ91QU	Vontobel	12025	0.43	11827.12	7	10000	7.7977
AP6MBQ	Interactive Brokers	12200	0.13	11833.37	7	10000	7.7402
PS2SLE	BNP Paribas	11950	0.66	11830.66	7	10000	7.7262
GL15K1	Goldman Sachs	12200	0.122	11826.48	7	10000	7.7042
VZ91QT	Vontobel	11975	0.58	11830.46	7	10000	7.6936
CN263S	Commerzbank	12175	0.16	11831.17	7	10000	7.6734
UZ7LUV	UBS	11750	1.42	11827.42	5	10000	7.6701
VZ91QY	Vontobel	12225	0.1	11831.07	7	10000	7.6400
CC8Z2L	Citi	12050	0.38	11829.59	7	10000	7.6101
XM2HAN	Deutsche Bank	12150	0.11	11833.11	5	10000	7.5726
CN263N	Commerzbank	11975	0.58	11829.5	7	10000	7.5534
VZ9CZ1	Vontobel	11950	0.67	11831.82	7	10000	7.4668

A.3 Source codes

In this section we present the source codes used to generate the plots in Chapter 7. Main parts of the source codes are modified versions of the code which is given in [120]. Firstly, we calculate M paths of the forward process $S^N_{t_k}$ with initial value $S_0 = 100$ for a fixed seed:

Listing A.1: Forward process.

```
% fix seed
rng(seed);

[X,W] = ForwardProcess(n+1,M,T,100);

% Forward process:
%
% This function simulates M paths of the forward process X^N_tk
% using Euler scheme starting at initial value X0 with constant
% drift mu and standard deviation sigma.

function [X,dW] = ForwardProcess(N,M,T,X0,mu,sigma)

h = T/N;
X = zeros(N,M);

```

```
NormInc = randn(N,M);
NormInc(1,:) = zeros(1,M);
dW = sqrt(h)*NormInc;

X(1,:) = X0;

    for m = 2:N
        X(m,:) = X(m-1,:) + mu*X(m-1,:)*h
                                      + sigma*X(m-1,:).*dW(m,:);
    end % for

end % function
```

The results in Section 7.4.2 have been computed with the following source codes:

Listing A.2: Source codes for Romanovski's approach.

```
% Composed RM:
%
% The following code approximates composed V@R and AV@R along the
% simulated underlying asset paths.
%
% Input arguments are R (polynomial degree), M (number of Monte
% Carlo paths of forward process), K (number of one time step long
% Monte Carlo simulations; see function ForwardStepMat), N
% (number of time steps), T (terminal time), aT (level of (A)V@R)
% and the options polBase ('lag' (laguerre polynomial base) or
% 'her' (hermite polynomial base)) and opt ('var' (V@R) or
% 'avar' (AV@R)). We also pass the result X from the forward
% process.
%
% The algorithm returns an estimate of the risk measure at the
% initial time.

function est = ComMeasure(R,M,K,N,T,aT,polBase,opt,X)

est = zeros(3,1);

YT(:,:,1) = Phi(X(N+1,:),'call');
YT(:,:,2) = Phi(X(N+1,:),'barrier');
YT(:,:,3) = Phi(X(N+1,:),'smooth');


h = T/N;
Y = zeros(N+1,M,3);
Alpha = zeros(N,R+1,3);
Y(N+1,:,1) = YT(:,:,1);
```

```
Y(N+1,:,1) = YT(:,:,2);
Y(N+1,:,1) = YT(:,:,3);

switch opt

    case 'var'
        alpha = 1-normcdf(norminv(1-aT,0,1)*sqrt(h/T));

    case 'avar'
        no = 14; % accuracy of the calculation of alpha_h

        alpha = alp(N,T,aT,no);

end % switch

l = 2;
r = 0.5;

for m = N:-1:1

    Paths=ForwardStepMat(K,h,X(m,:));
    Values(:,:,1) = Paths.*0;
    Values(:,:,2) = Paths.*0;
    Values(:,:,3) = Paths.*0;

    switch opt

        case 'var'
            if m == N
                Values(:,:,1) = Phi(Paths,'call');
                Values(:,:,2) = Phi(Paths,'barrier');
                Values(:,:,3) = Phi(Paths,'smooth');

            else
                for i = 1:M
                    pComRM = base(R,Paths(:,i)',l*((m+2)/T)^r
                                                       ,polBase)';
                     Values(:,i,1) = (Alpha(m+1,:,1)*pComRM')';
                     Values(:,i,2) = (Alpha(m+1,:,2)*pComRM')';
                     Values(:,i,3) = (Alpha(m+1,:,3)*pComRM')';
                end % for
            end % if
                Y(m,:,1) = quantile(Values(:,:,1),1-alpha);
                Y(m,:,2) = quantile(Values(:,:,2),1-alpha);
                Y(m,:,3) = quantile(Values(:,:,3),1-alpha);
```

```

        case 'avar'
            if m == N
                Values(:,:,1) = Phi(Paths,'call');
                Values(:,:,2) = Phi(Paths,'barrier');
                Values(:,:,3) = Phi(Paths,'smooth');

                quantiles(:,1) = quantile(Values(:,:,1),1-alpha);
                quantiles(:,2) = quantile(Values(:,:,2),1-alpha);
                quantiles(:,3) = quantile(Values(:,:,3),1-alpha);

                es = zeros(M,3);
                for i = 1:M
                    tmp = Values(Values(:,i,1)>=quantiles(i,1)
                                                            ,i,1);
                    es(i,1) = sum(tmp)/length(tmp);
                    tmp = Values(Values(:,i,2)>=quantiles(i,2)
                                                            ,i,2);
                    es(i,2) = sum(tmp)/length(tmp);
                    tmp = Values(Values(:,i,3)>=quantiles(i,3)
                                                            ,i,3);
                    es(i,3) = sum(tmp)/length(tmp);
                end % parfor
            else
                a = Alpha(m+1,:,:);
                es = zeros(M,3);
                for i = 1:M
                    pComRM = base(R,Paths(:,i)',l*((m+2)/T)^r
                                                     ,polBase)';
                    vals(:,:,1) = (a(:,:,1)*pComRM')';
                    VaR(:,1) = quantile(vals(:,:,1),1-alpha);
                    tmp = vals(vals(:,:,1) >=VaR(:,1),:,1);
                    es(i,1) = sum(tmp)/length(tmp);

                    vals(:,:,2) = (a(:,:,2)*pComRM')';
                    VaR(:,2) = quantile(vals(:,:,2),1-alpha);
                    tmp = vals(vals(:,:,2) >=VaR(:,2),:,2);
                    es(i,2) = sum(tmp)/length(tmp);

                    vals(:,:,3) = (a(:,:,3)*pComRM')';
                    VaR(:,3) = quantile(vals(:,:,3),1-alpha);
                    tmp = vals(vals(:,:,3) >=VaR(:,3),:,3);
                    es(i,3) = sum(tmp)/length(tmp);
                end % for
            end % if
```

```
                Y(m,:,1) = es(:,1);
                Y(m,:,2) = es(:,2);
                Y(m,:,3) = es(:,3);
    end % switch

        A = base(R,X(m,:),l*((m+1)/T)^r,polBase)';
        Alpha(m,:,1) = (A\Y(m,:,1)')';
        Y(m,:,1) = (Alpha(m,:,1)*A');
        Alpha(m,:,2) = (A\Y(m,:,2)')';
        Y(m,:,2) = (Alpha(m,:,2)*A');
        Alpha(m,:,3) = (A\Y(m,:,3)')';
        Y(m,:,3) = (Alpha(m,:,3)*A');

end % for

est(1) = Y(1,1,1);
est(2) = Y(1,1,2);
est(3) = Y(1,1,3);

end % function


% BSDE solver:
%
% The following code approximates the BSDE solution using the
% regression based MC method
%
% Input arguments are R (polynomial degree), M (number of Monte
% Carlo paths of forward process), N (number of time steps),
% T (terminal time), the options polBase ('lag' (laguerre
% polynomial base) or 'her' (hermite polynomial base)) and opt
% ('var' (V@R) or 'avar' (AV@R)) and aT (level of (A)V@R).
% We also pass the results X and W from the forward process.
%
% The algorithm returns the estimate of Y^N_0.

function est = BSDESolver(R,M,N,T,polBase,opt,X,W,aT)

est = zeros(3,1);
Y = zeros(N+1,M,3);
Z = zeros(N,M,3);
Beta = zeros(N,R+1,3);
Alpha = zeros(N,R+1,3);

YT(:,:,1) = Phi(X(N+1,:),'call');
```

```
YT(:,:,2) = Phi(X(N+1,:),'barrier');
YT(:,:,3) = Phi(X(N+1,:),'smooth');

h = T/N;
l = 2;
r = 0.5;
I = 3;

Y(N+1,:,1) = YT(:,:,1);
Y(N+1,:,2) = YT(:,:,2);
Y(N+1,:,3) = YT(:,:,3);

     for m = N:-1:1

          pBSDE = base(R,X(m,:), l*((m+1)/T)^r, polBase)';

          for o = 1:3

             for i = 1:I;
                  b(1,:,o) = Y(m+1,:,o)+h*F(Beta(m,:,o)*pBSDE',opt
                                                                   ,h,aT);
                  A = [pBSDE, repmat(W(m+1,:)',1,R+1).*pBSDE];
                  coeff(:,o) = A\(b(1,:,o))';
                  Alpha(m,:,o) = coeff(1:R+1,o);
                  Beta(m,:,o) = coeff(R+2:end,o);
             end % for

             Y(m,:,o) = (Alpha(m,:,o)*pBSDE');
             Z(m,:,o) = (Beta(m,:,o)*pBSDE');

          end

     end % for

est(1) = Y(1,1,1);
est(2) = Y(1,1,2);
est(3) = Y(1,1,3);

end % function




%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Additional Functions:
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%


% Simulating paths from each point:
%
% This functions performs K one time step long Monte Carlo
% simulations of forward process X starting at X0, where X0 can
% be a vector, in which case function ForwardStepMat performs
% K simulations from each point in X0.
%
% A matrix of simulated values is returned with each column
% representing paths from one of those points in X0. Other
% required input data is the size of the time step h.

function X = ForwardStepMat(K,h,X0,mu,sigma)

X = repmat(X0,K,1);

NormInc = randn(K,length(X0));
dW = sqrt(h)*NormInc;
muVal = repmat(mu*X0,K,1);
sigmaVal = repmat(sigma*X0,K,1);

X = X + muVal*h + sigmaVal.*dW;

end % function


% BSDE terminal function:

function [outphi] = Phi(XT,payoff)
[a,b] = size(XT);
outphi = zeros(a,b);
    for i = 1:a
        for j = 1:b

        switch payoff
            case 'call' % call-option
                outphi(i,j) = max(XT(i,j)-100,0);
            case 'barrier'% barrier-payoff
                if XT(i,j) > =135
                    outphi(i,j) = 70;
                end % if
            case 'smooth'% smooth payoff
                outphi(i,j) = log(XT(i,j)+1);
```

```
        end % switch

        end % for
    end % for

end % function


% Evaluating polynomial basis:
%
% The following code returns a k + 1 by length(x) matrix with the
% first k base polynomials evaluated at x/d, multiplied by
% exponential weighting. It supports two different polynomial
% bases, Laguerre and Hermite polynomial bases, through the
% choice of parameter opt, accepting arguments lag (Laguerre
% polynomial base) or her (Hermite polynomial base).

function [ L ] = base(k,x,d,opt)

x = x/d;
L = zeros([k+1,length(x)]);
L(1,:) = 1;
w = exp(-0.5*abs(x));

    switch opt

        case 'lag'
            if k >= 2
                L(2,:) = (1-x);
            end % if

            for j = 3:(k+1)
                L(j,:)=w.*((2*(j-2)+1-x).*L(j-1,:)-(j-2)*L(j-2,:))
                                                                /(j-1);
            end % for

        case 'her'
            if k >= 2
                L(2,:) = x;
            end % if

            for j = 3:(k+1)
                L(j,:) = w.*(x.*L(j-1,:)-(j-1)*L(j-2,:));
            end % for

```

```
    end % switch

end % function

% BSDE driver:
%
% Given a specific value of parameter opt a driver for the BSDE
% is returned, solution of which is the driver for continuous
% V@R, or the driver for continuous AV@R.

function [outF] = F(z,opt,T,aT)

switch opt

        case 'var'
            outF = norminv(1-aT,0,1)*abs(z)/sqrt(T);

        case 'avar'
            outF = exp(-norminv(1-aT,0,1)^2/2)/(sqrt(2*pi*T)*(aT))
                                                            *abs(z);

end % switch

end % function

% Function that computes alpha for AV@R

function erg = alp(N,T,aT,n)

if mod(n,2) == 0
    n=n+1;
end

eta = 10^-n;
h = T/N;
syms z;
nmlinv(z) = -sqrt(2)*erfcinv(2*z); % Equivalent to 'norminv'

k=(1-eta);
ergebnis = vpa(1/((k)*sqrt(2*pi*h))*exp(-nmlinv(1-k)^2/2)
                    -1/((aT)*sqrt(2*pi*T))*exp(-nmlinv(1-aT)^2/2));

for i = n:-1:1

    if mod(i,2) == 1
```

```

            while (ergebnis < 0)
                k = k-eta*10^(i-1);
              ergebnis=vpa(1/((k)*sqrt(2*pi*h))*exp(-nmlinv(1-k)^2/2)
                      -1/((aT)*sqrt(2*pi*T))*exp(-nmlinv(1-aT)^2/2));
            end

        else

            while (ergebnis > 0)
                k = k+eta*10^(i-1);
              ergebnis=vpa(1/((k)*sqrt(2*pi*h))*exp(-nmlinv(1-k)^2/2)
                      -1/((aT)*sqrt(2*pi*T))*exp(-nmlinv(1-aT)^2/2));
            end

        end

end

erg = k;

end
```

To compute the results in Section 7.4.2 we have used the following source codes:

Listing A.3: Source codes for Stadje's approach.

```
% Composed scaled RM:
%
% The following code approximates composed V@R or AV@R along the
% simulated underlying asset paths following Stadjes approach.
%
% Input arguments are R (polynomial degree), M (number of Monte
% Carlo paths of forward process), K (number of one time step
% long Monte Carlo simulations; see function ForwardStepMat), N
% (number of time steps), T (terminal time), aT (level of (A)V@R)
% and the options polBase ('lag' (laguerre polynomial base) or
% 'her' (hermite polynomial base)) and opt ('var' (V@R) or 'avar'
% (AV@R)). We also pass the result X from the forward process.
%
% The algorithm returns an estimate of the risk measure at the
% initial time.

function est = ComMeasure(R,M,K,N,T,aT,polBase,opt,X)

est=zeros(3,1);

```

```
YT(:,:,1) = Phi(X(N+1,:),'call');
YT(:,:,2) = Phi(X(N+1,:),'barrier');
YT(:,:,3) = Phi(X(N+1,:),'smooth');


h = T/N;
Y = zeros(N+1,M,3);
Alpha = zeros(N,R+1,3);
Y(N+1,:,1) = YT(:,:,1);
Y(N+1,:,1) = YT(:,:,2);
Y(N+1,:,1) = YT(:,:,3);

switch opt

    case 'var'
        alpha = aT;

    case 'avar'
        alpha = aT;

end % switch

l = 2;
r = 0.5;

for m = N:-1:1

    Paths = ForwardStepMat(K,h,X(m,:));
    Values(:,:,1) = Paths.*0;
    Values(:,:,2) = Paths.*0;
    Values(:,:,3) = Paths.*0;

    switch opt

        case 'var'
            if m == N
                Values(:,:,1) = Phi(Paths,'call');
                Values(:,:,2) = Phi(Paths,'barrier');
                Values(:,:,3) = Phi(Paths,'smooth');

            else
                for i=1:M
                    pComRM = base(R,Paths(:,i)',l*((m+2)/T)^r
                                                      ,polBase)';
                     Values(:,i,1) = (Alpha(m+1,:,1)*pComRM')';
```

```
                    Values(:,i,2) = (Alpha(m+1,:,2)*pComRM')';
                    Values(:,i,3) = (Alpha(m+1,:,3)*pComRM')';
                end % for
            end % if
                for i=1:M
                    Y(m,i,1) = (1/K)*sum(Values(:,i,1))+sqrt(h)
        *quantile(Values(:,i,1)-(1/K)*sum(Values(:,i,1)),1-alpha);
                    Y(m,i,2) = (1/K)*sum(Values(:,i,2))+sqrt(h)
        *quantile(Values(:,i,2)-(1/K)*sum(Values(:,i,2)),1-alpha);
                    Y(m,i,3) = (1/K)*sum(Values(:,i,3))+sqrt(h)
        *quantile(Values(:,i,3)-(1/K)*sum(Values(:,i,3)),1-alpha);
                end % for

        case 'avar'
            if m == N
                Values(:,:,1) = Phi(Paths,'call');
                Values(:,:,2) = Phi(Paths,'barrier');
                Values(:,:,3) = Phi(Paths,'smooth');

                for i=1:M
                quantiles(i,1) = quantile(Values(:,i,1)
                                    -(1/K)*sum(Values(:,i,1)),1-alpha);
                quantiles(i,2) = quantile(Values(:,i,2)
                                    -(1/K)*sum(Values(:,i,2)),1-alpha);
                quantiles(i,3) = quantile(Values(:,i,3)
                                    -(1/K)*sum(Values(:,i,3)),1-alpha);
                end % for

                es = zeros(M,3);
                for i = 1:M
                    tmp=Values(Values(:,i,1)
                   -(1/K)*sum(Values(:,i,1))>=quantiles(i,1),i,1);
                    es(i,1) = sum(tmp-(1/K)*sum(Values(:,i,1)))
                                                        /length(tmp);
                    comes(i,1) = (1/K)*sum(Values(:,i,1))+sqrt(h)
                                                            *es(i,1);

                    tmp=Values(Values(:,i,2)
                   -(1/K)*sum(Values(:,i,2))>=quantiles(i,2),i,2);
                    es(i,2) = sum(tmp-(1/K)*sum(Values(:,i,2)))
                                                        /length(tmp);
                    comes(i,2) = (1/K)*sum(Values(:,i,2))+sqrt(h)
                                                            *es(i,2);

                    tmp=Values(Values(:,i,3)
```

```
                        -(1/K)*sum(Values(:,i,3))>=quantiles(i,3),i,3);
                    es(i,3) = sum(tmp-(1/K)*sum(Values(:,i,3)))
                                                         /length(tmp);
                    comes(i,3) = (1/K)*sum(Values(:,i,3))+sqrt(h)
                                                             *es(i,3);
                end % for
            else
                a = Alpha(m+1,:,:);
                es = zeros(M,3);
                for i = 1:M
                    pComRM = base(R,Paths(:,i)',l*((m+2)/T)^r
                                                  ,polBase)';
                    vals(:,:,1) = (a(:,:,1)*pComRM')';
                    VaR(:,1) = quantile(vals(:,:,1)
                                  -(1/K)*sum(vals(:,:,1)),1-alpha);
                    tmp = vals(vals(:,:,1)-(1/K)*sum(vals(:,:,1))
                                                  >=VaR(:,1),:,1);
                    es(i,1) = sum(tmp-(1/K)*sum(vals(:,:,1)))
                                                         /length(tmp);
                    comes(i,1)=(1/K)*sum(vals(:,:,1))+sqrt(h)
                                                             *es(i,1);

                    vals(:,:,2) = (a(:,:,2)*pComRM')';
                    VaR(:,2) = quantile(vals(:,:,2)
                                  -(1/K)*sum(vals(:,:,2)),1-alpha);
                    tmp = vals(vals(:,:,2)-(1/K)*sum(vals(:,:,2))
                                                  >=VaR(:,2),:,2);
                    es(i,2) = sum(tmp-(1/K)*sum(vals(:,:,2)))
                                                         /length(tmp);
                    comes(i,2)=(1/K)*sum(vals(:,:,2))+sqrt(h)
                                                             *es(i,2);

                    vals(:,:,3) = (a(:,:,3)*pComRM')';
                    VaR(:,3) = quantile(vals(:,:,3)
                                  -(1/K)*sum(vals(:,:,3)),1-alpha);
                    tmp = vals(vals(:,:,3)-(1/K)*sum(vals(:,:,3))
                                                  >=VaR(:,3),:,3);
                    es(i,3) = sum(tmp-(1/K)*sum(vals(:,:,3)))
                                                         /length(tmp);
                    comes(i,3)=(1/K)*sum(vals(:,:,3))+sqrt(h)
                                                             *es(i,3);
                end % for
            end % if
                Y(m,:,1) = comes(:,1);
                Y(m,:,2) = comes(:,2);
```

```
                Y(m,:,3) = comes(:,3);
    end % switch

        A = base(R,X(m,:),l*((m+1)/T)^r,polBase)';

        Alpha(m,:,1)=(A\Y(m,:,1)')';
        Y(m,:,1)=(Alpha(m,:,1)*A');
        Alpha(m,:,2)=(A\Y(m,:,2)')';
        Y(m,:,2)=(Alpha(m,:,2)*A');
        Alpha(m,:,3)=(A\Y(m,:,3)')';
        Y(m,:,3)=(Alpha(m,:,3)*A');

end % for

est(1) = Y(1,1,1);
est(2) = Y(1,1,2);
est(3) = Y(1,1,3);

end % function


% BSDE solver:
%
% The following code approximates the BSDE solution using the
% regression based MC method
%
% Input arguments are R (polynomial degree), M (number of Monte
% Carlo paths of forward process), N (number of time steps),
% T (terminal time), the options polBase ('lag' (laguerre
% polynomial base) or 'her' (hermite polynomial base)) and opt
% ('var' (V@R) or 'avar' (AV@R)) and aT (level of (A)V@R).
% We also pass the results X and W from the forward process.
%
% The algorithm returns the estimate of Y^N_0.

function est = BSDESolver(R,M,N,T,polBase,opt,X,W,aT)

est = zeros(3,1);
Y = zeros(N+1,M,3);
Z = zeros(N,M,3);
Beta = zeros(N,R+1,3);
Alpha = zeros(N,R+1,3);

YT(:,:,1) = Phi(X(N+1,:),'call');
YT(:,:,2) = Phi(X(N+1,:),'barrier');
```

```
YT(:,:,3) = Phi(X(N+1,:),'smooth');

h = T/N;
l = 2;
r = 0.5;
I = 3;

Y(N+1,:,1) = YT(:,:,1);
Y(N+1,:,2) = YT(:,:,2);
Y(N+1,:,3) = YT(:,:,3);

    for m = N:-1:1

        pBSDE = base(R,X(m,:),  l*((m+1)/T)^r,polBase)';

        for o = 1:3

          for i = 1:I;
              b(1,:,o) = Y(m+1,:,o) + h*F(Beta(m,:,o)*pBSDE',opt
                                                              ,aT);
              A = [pBSDE,repmat(W(m+1,:)',1,R+1).*pBSDE];
              coeff(:,o) = A\(b(1,:,o))';
              Alpha(m,:,o) = coeff(1:R+1,o);
              Beta(m,:,o) = coeff(R+2:end,o);
          end % for

          Y(m,:,o) = (Alpha(m,:,o)*pBSDE');
          Z(m,:,o) = (Beta(m,:,o)*pBSDE');

        end

    end % for

est(1) = Y(1,1,1);
est(2) = Y(1,1,2);
est(3) = Y(1,1,3);

end % function




%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Additional Functions:
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
% BSDE driver:
%
% Given a specific value of parameter opt a driver for the BSDE
% is returned, solution of which is the driver for continuous V@R,
% or the driver for continuous AV@R.
%
% In contrast to the BSDE driver function used in the approach of
% Romanovski, this function is independent of T.

function [outF] = F(z,opt,aT)

    switch opt

        case 'var'
            outF = norminv(1-aT,0,1)*abs(z);

        case 'avar'
            outF = exp(-norminv(1-aT,0,1)^2/2)/(sqrt(2*pi)*aT)
                                                          *abs(z);

    end % switch

end % function
```

Scientific Career

06/2006	Abitur at Albert-Einstein-Gymnasium Völklingen
11/2008 - 08/2010	Student assistant at the Faculty of Mathematics at the University of Kaiserslautern
10/2012	Diploma in Management Mathematics (Dipl.-Math. oec.) (Specialization: Financial Mathematics)
01/2013 - 12/2014	Scientific assistant at the Chair of Macroeconomics at the University of Kaiserslautern
11/2012 - 10/2015	Ph.D. student of Prof. Dr. Jörn Saß at the University of Kaiserslautern

Wissenschaftlicher Werdegang

06/2006	Abitur am Albert-Einstein-Gymnasium in Völklingen
11/2008 - 08/2010	Studentische Hilfskraft am Fachbereich Mathematik an der TU Kaiserslautern
10/2012	Diplom in Wirtschaftsmathematik (Dipl.-Math. oec.) (Studienschwerpunkt: Finanzmathematik)
01/2013 - 12/2014	Wissenschaftllicher Mitarbeiter am Lehrstuhl Makroökonomik an der TU Kaiserslautern
11/2012 - 10/2015	Doktorand von Prof. Dr. Jörn Saß an der TU Kaiserslautern