Finance Stochast. 6, 429–447 (2002)

Finance and Stochastics

Convex measures of risk and trading constraints

Hans Föllmer, Alexander Schied

Institut für Mathematik, Humboldt-Universität, Unter den Linden 6, 10099 Berlin, Germany
(e-mail: {foellmer,schied}@mathematik.hu-berlin.de)

Abstract. We introduce the notion of a convex measure of risk, an extension of the concept of a coherent risk measure defined in Artzner et al. (1999), and we prove a corresponding extension of the representation theorem in terms of probability measures on the underlying space of scenarios. As a case study, we consider convex measures of risk defined in terms of a robust notion of bounded shortfall risk. In the context of a financial market model, it turns out that the representation theorem is closely related to the superhedging duality under convex constraints.

Key words: Risk measure, convex measure of risk, shortfall, trading constraints, efficient hedging

JEL Classification: G18, G11, C60, C61

Mathematics Subject Classification (2000): 46N10, 91B16, 91B28, 91B30

1 Introduction

There is a considerable interest, both from a theoretical and a practical point of view, in a quantitative assessment of the risk involved in a financial position. If such a position is described by the resulting discounted net worth at the end of a given period, defined as a real-valued function X on some set Ω of possible scenarios, then a quantitative *measure of risk* is given by a mapping ρ from a certain space $\mathcal{X}$ of functions on Ω to the real line. Clearly, such a map ρ should satisfy certain conditions of consistency.

Manuscript received: December 2000; final version received: January 2002

Recently, Artzner et al. [1] have introduced the concept of a *coherent measure of risk*. It is defined by the following properties of the mapping $\rho : \mathcal{X} \to \mathbb{R}$:

$$\text{Subadditivity: } \rho(X+Y) \le \rho(X) + \rho(Y). \tag{1}$$

$$\text{Positive Homogeneity: If } \lambda \ge 0, \text{ then } \rho(\lambda X) = \lambda \rho(X). \tag{2}$$

$$\text{Monotonicity: If } X \le Y, \text{ then } \rho(X) \ge \rho(Y). \tag{3}$$

$$\text{Translation Invariance: If } m \in \mathbb{R}, \text{ then } \rho\big(Y+m\big) = \rho(Y) - m. \tag{4}$$

Typically, a coherent measure of risk ρ arises from some family $\mathcal{Q}$ of probability measures on Ω by computing the expected loss under $Q \in \mathcal{Q}$ and then taking the worst result as Q varies over $\mathcal{Q}$:

$$\rho(X) = \sup_{Q \in \mathcal{Q}} E_Q[\,-X\,]; \tag{5}$$

for the case where Ω is a finite set see Artzner et al. [1] and, in a different context, Huber [11].

In many situations, however, the risk of a position might increase in a nonlinear way with the size of the position. For example, an additional liquidity risk may arise if a position is multiplied by a large factor. This suggests to relax the conditions of positive homogeneity and of subadditivity and to require, instead of (1) and (2), the weaker property of

$$\text{Convexity: } \rho\big(\lambda X + (1-\lambda)Y\big) \le \lambda\rho(X) + (1-\lambda)\rho(Y) \text{ for any } \lambda \in [0,1]. \tag{6}$$

Convexity means that diversification does not increase the risk, i.e., the risk of a diversified position $\lambda X + (1-\lambda)Y$ is less or equal to the weighted average of the individual risks. Let $\mathcal{X}$ be a convex set of functions on the set Ω of possible scenarios. We assume that $0 \in \mathcal{X}$ and that $\mathcal{X}$ is closed under the addition of constants.

Definition 1 *A map $\rho : \mathcal{X} \to \mathbb{R}$ will be called a convex measure of risk if it satisfies the condition of convexity (6), monotonicty (3), and translation invariance (4).*

If the convex measure of risk ρ is normalized in the sense that $\rho(0) = 0$, then the quantity $\rho(X)$ can be interpreted as a "margin requirement", i.e., the minimal amount of capital which, if added to the position at the beginning of the given period and invested into a risk-free asset, makes the discounted position X "acceptable".

In Sect. 2 we prove a representation theorem for convex measures of risk. In the case in which $\mathcal{X}$ is the space of all real-valued functions on a finite set Ω, any convex measure of risk is of the form

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

where $\mathcal{P}$ is the set of all probability measures on Ω, and $\alpha(\cdot)$ is a certain "penalty function" on $\mathcal{P}$. Independently, this result was stated by D. Heath in [9]; see also [10]. In Theorem 6, we prove an appropriate extension of the representation theorem

where the finite set Ω is replaced by a general probability space $(\Omega, \mathcal{F}, P)$, and $\mathcal{X}$ is the space $L^\infty(\Omega, \mathcal{F}, P)$ of bounded random variables.

An alternative representation in terms of probability measures, which does not use a given reference measure P, is discussed in [15].

A convex measure of risk can be characterized in terms of properties of the associated *acceptance set*

$$\mathcal{A}_\rho = \left\{ X \in \mathcal{X} \,\middle|\, \rho(X) \le 0 \right\}.$$

Conversely, a given set $\mathcal{A}$ of "acceptable positions" defines a convex measure of risk via

$$\rho_{\mathcal{A}}(X) := \inf \left\{ m \in \mathbb{R} \,\middle|\, m + X \in \mathcal{A} \right\}$$

provided that $\mathcal{A}$ satisfies itself certain axioms. Our two main case studies of convex risk measures will be defined in terms of such acceptance sets.

In Sect. 3, we consider the situation where the acceptance set is defined in terms of a (robust) notion of bounded shortfall risk, i.e., a position is acceptable if its shortfall risk is bounded by some given level. In this case, the associated penalty function can be described by a functional of the type of a dual Orlicz norm.

In Sect. 4, we consider a discrete-time financial market model with convex trading constraints. As a first step, we define a set $\mathcal{A}$ of acceptable positions as the class of functions $X \in \mathcal{X}$ such that $X + V \ge 0$ where V is the final portfolio value which can be generated by an admissible trading strategy. Theorem 16 clarifies the structure of the corresponding risk measure $\rho_{\mathcal{A}}$. In this case, the weight function $\alpha(Q)$ can be described explicitly in terms of the increasing process which appears in the construction of the optional decomposition under convex trading constraints in Föllmer and Kramkov [5].

2 Convex measures of risks

2.1 Acceptance sets

Let $\mathcal{X}$ be a linear space of functions on a given set Ω of possible scenarios. We assume that $\mathcal{X}$ contains all constant functions. Any risk measure $\rho : \mathcal{X} \to \mathbb{R}$ induces an *acceptance set* $\mathcal{A}_\rho$ defined as

$$\mathcal{A}_\rho := \left\{ X \in \mathcal{X} \,\middle|\, \rho(X) \le 0 \right\}. \tag{7}$$

Conversely, for a given class $\mathcal{A}$ of acceptable positions, we can introduce an associated risk measure $\rho_{\mathcal{A}}$ by defining

$$\rho_{\mathcal{A}}(X) := \inf \left\{ m \in \mathbb{R} \,\middle|\, m + X \in \mathcal{A} \right\}. \tag{8}$$

The following two propositions summarize the relations between a convex measure of risk and its acceptance set $\mathcal{A}_\rho$. They are similar to the ones found for coherent measures of risk; cf. [1], [4], [12].

Proposition 2 *Suppose $\rho : \mathcal{X} \to \mathbb{R}$ is a convex measure of risk with associated acceptance set $\mathcal{A}_\rho$. Then*

$$\rho_{\mathcal{A}_\rho} = \rho.$$

Moreover, $\mathcal{A} := \mathcal{A}_\rho$ enjoys the following properties.

1. *$\mathcal{A}$ is convex and non-empty.*
2. *If $X \in \mathcal{A}$ and $Y \in \mathcal{X}$ satisfies $Y \geq X$, then $Y \in \mathcal{A}$.*
3. *If $X \in \mathcal{A}$ and $Y \in \mathcal{X}$, then*

$$\big\{\, \lambda \in [0,1] \,\big|\, \lambda X + (1-\lambda)Y \in \mathcal{A} \,\big\}$$

is closed in $[0,1]$.

Proof To show that $\rho_{\mathcal{A}_\rho}(X) = \rho(X)$ for all X, note that the translation invariance of ρ implies that

$$\begin{aligned}\rho_{\mathcal{A}_\rho}(X) &= \inf\big\{\, m \,\big|\, m + X \in \mathcal{A}_\rho \,\big\} = \inf\big\{\, m \,\big|\, \rho(m+X) \leq 0 \,\big\} \\ &= \inf\big\{\, m \,\big|\, \rho(X) \leq m \,\big\} = \rho(X)\,.\end{aligned}$$

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

Example 3 (Value at Risk) Value at Risk at level $\gamma > 0$,

$$\mathit{VaR}_\gamma(X) := \inf\Big\{\, m \,\Big|\, P\big[\, X + m < 0 \,\big] \leq \gamma \,\Big\}\,,$$

is *not* a convex measure of risk. This is shown by the example in [1], p. 218, since the acceptance set is not convex.

Proposition 4 *Assume that $\mathcal{A} \neq \emptyset$ is a convex subset of $\mathcal{X}$ which satisfies property 2 of Proposition 2, and denote by $\rho_{\mathcal{A}}$ the functional associated to $\mathcal{A}$ via (8). If $\rho_{\mathcal{A}}(0) > -\infty$, then*

1. *$\rho_{\mathcal{A}}$ is a convex measure of risk.*
2. *$\mathcal{A}$ is a subset of $\mathcal{A}_{\rho_{\mathcal{A}}}$. Moreover, if $\mathcal{A}$ satisfies property 3 of Proposition 2, then $\mathcal{A} = \mathcal{A}_{\rho_{\mathcal{A}}}$.*

Proof 1. It is straightforward to verify that $\rho_{\mathcal{A}}$ satisfies translation invariance and monotonicity. We show next that $\rho_{\mathcal{A}}$ takes only finite values. To this end, fix some element Y of the non-empty set $\mathcal{A}$. For $X \in \mathcal{X}$ given, there exists a finite number m with $m + X > Y$, because X and Y are both bounded. Monotonicity, translation invariance, and the fact that $\rho_{\mathcal{A}}(Y) \leq 0$ give $\rho_{\mathcal{A}}(X) \leq m$. To show that $\rho_{\mathcal{A}}(X) > -\infty$, we take m' such that $X + m' \leq 0$ and conclude that $\rho_{\mathcal{A}}(X) \geq \rho_{\mathcal{A}}(0) + m' > -\infty$.

As to the property of convexity, suppose that $X_1,\ X_2 \in \mathcal{X}$ and that $m_1,\ m_2 \in \mathbb{R}$ are such that $m_i + X_i \in \mathcal{A}$. If $\lambda \in [0,1]$, then the convexity of $\mathcal{A}$ implies that $\lambda(m_1 + X_1) + (1-\lambda)(m_2 + X_2) \in \mathcal{A}$. Thus, by the translation invariance of $\rho_{\mathcal{A}}$,

$$\begin{aligned} 0 &\geq \rho_{\mathcal{A}}\big(\lambda(m_1 + X_1) + (1-\lambda)(m_2 + X_2)\big) \\ &= \rho_{\mathcal{A}}\big(\lambda X_1 + (1-\lambda)X_2\big) - \big(\lambda m_1 + (1-\lambda)m_2\big)\,, \end{aligned}$$

and the convexity of $\rho_{\mathcal{A}}$ follows.

2. The inclusion $\mathcal{A} \subseteq \mathcal{A}_{\rho_{\mathcal{A}}}$ is obvious. Now assume that $\mathcal{A}$ satisfies the third property of Proposition 2. We have to show that $X \notin \mathcal{A}$ implies that $\rho_{\mathcal{A}}(X) > 0$. To this end, take $m > \rho_{\mathcal{A}}(0)$. By property 3 of Proposition 2, there exists an $\varepsilon \in (0,1)$ such that $\varepsilon m + (1-\varepsilon)X \notin \mathcal{A}$. Thus,

$$\begin{aligned} \varepsilon m \leq \rho_{\mathcal{A}}\big((1-\varepsilon)X\big) &= \rho_{\mathcal{A}}\big(\varepsilon \cdot 0 + (1-\varepsilon)X\big) \\ &\leq \varepsilon\rho_{\mathcal{A}}(0) + (1-\varepsilon)\rho_{\mathcal{A}}(X)\,. \end{aligned}$$

Hence

$$\rho_{\mathcal{A}}(X) \geq \frac{\varepsilon\big(m - \rho_{\mathcal{A}}(0)\big)}{1-\varepsilon} > 0\,,$$

and property 2 follows. □

2.2 The representation theorem for convex measures of risk

Now we prove the structure theorem for convex measure of risks. Let us first consider the special case in which $\mathcal{X}$ is the space of all real-valued functions on some finite set Ω.

Theorem 5 *Suppose $\mathcal{X}$ is the space of all real-valued functions on a finite set Ω. Then $\rho : \mathcal{X} \to \mathbb{R}$ is a convex measure of risk if and only if there exists a "penalty function" $\alpha : \mathcal{P} \to (-\infty, \infty]$ such that*

$$\rho(Z) = \sup_{Q \in \mathcal{P}} \Big(E_Q\big[-Z\big] - \alpha(Q) \Big)\,. \tag{10}$$

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

Note that this theorem includes the structure theorem for coherent measures of risk as a special case. Indeed, it is easy to see that ρ will possess the property of positive homogeneity, i.e., ρ will be a coherent measure of risk, if and only if the particular penalty function $\alpha(\cdot)$ constructed in the proof takes only the values 0 and $+\infty$. In this case, our theorem implies the representation (5) in terms of the set

$$\mathcal{Q} = \Big\{\, Q \in \mathcal{P} \,\Big|\, \alpha(Q) = 0 \,\Big\}\,.$$

Proof of Theorem 5 The "if"-part is straightforward: For each $Q \in \mathcal{P}$ the functional

$$X \mapsto E_Q\big[\,-X\,\big] - \alpha(Q)$$

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

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

$$\alpha(Q) := \sup_{X \in \mathcal{X}} \Big(E_Q\big[\,-X\,\big] - \rho(X) \Big)\,. \tag{11}$$

Then we claim that

$$\alpha(Q) = \sup_{X \in \mathcal{A}_\rho} E_Q\big[\,-X\,\big]\,. \tag{12}$$

For the moment denote the righthand side by $\widehat{\alpha}(Q)$. By definition of $\mathcal{A}_\rho$ we find $\alpha(Q) \geq \widehat{\alpha}(Q)$. To establish the converse inequality, take an arbitrary $X \in \mathcal{X}$ and recall that $X' := \rho(X) + X \in \mathcal{A}_\rho$. Thus

$$\widehat{\alpha}(Q) \geq E_Q\big[\,-X'\,\big] = E_Q\big[\,-X\,\big] - \rho(X)\,.$$

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

Now fix some $Y \in \mathcal{X}$ and take $\alpha(\cdot)$ as in (11). Then we clearly have

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

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

$$m > \sup_{Q \in \mathcal{P}} \Big(E_Q\big[\,-Y\,\big] - \alpha(Q) \Big)\,. \tag{13}$$

We must show that $m \geq \rho(Y)$ or, equivalently, $m + Y \in \mathcal{A}_\rho$. Suppose that, on the contrary, $m + Y \notin \mathcal{A}_\rho$. Since ρ is by definition a convex function on the Euclidean space $\mathbb{R}^\Omega$ taking only finite values, ρ is already continuous; cf. [14, Corollary 10.1.1]. Hence $A_\rho = \{\rho \leq 0\}$ is a closed convex set. Therefore, we can find a linear functional ℓ on $\mathbb{R}^\Omega$ such that

$$\beta := \sup_{X \in \mathcal{A}_\rho} \ell(X) < \ell\big(m + Y\big) =: \gamma < \infty\,. \tag{14}$$

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

$$\rho(X) \leq \rho(0) \text{ for } X \geq 0\,. \tag{15}$$

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

$$\gamma > \ell\big(\lambda X + \rho(0)\big) = \lambda \ell(X) + \rho(0)\,.$$

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

$$Q[\,A\,] := \ell\big(-\mathrm{I}_A\big)$$

defines a probability measure $Q \in \mathcal{P}$. By (12) and (14) we find

$$\alpha(Q) = \sup_{X \in \mathcal{A}_\rho} E_Q\big[-X\,\big] = \beta\,.$$

But

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

which is a contradiction to our choice of m. Therefore, we must have $m + Y \in \mathcal{A}_\rho$ and, thus, $m \geq \rho(Y)$. □

In the previous proof, the assumption that Ω is finite was only used in order to obtain the closedness of the acceptance set $\mathcal{A}_\rho$. In the case where $\mathcal{X}$ is given as the space $L^\infty(\Omega, \mathcal{F}, P)$ of bounded functions on a general probability space $(\Omega, \mathcal{F}, P)$, we will have to *assume* the closedness of $\mathcal{A}_\rho$ in a suitable topology, but then the previous argument goes through. Thus we obtain the following extension of Delbaen's representation theorem for coherent measures of risk on a general probability space; see [4, Theorem 3.2].

Theorem 6 *Suppose $\mathcal{X} = L^\infty(\Omega, \mathcal{F}, P)$, $\mathcal{P}$ is the set of probability measures $Q \ll P$, and $\rho : \mathcal{X} \to \mathbb{R}$ is a convex measure of risk. Then the following properties are equivalent.*

1. *There is a "penalty function" $\alpha : \mathcal{P} \to (-\infty, \infty]$ such that*

$$\rho(X) = \sup_{Q \ll P} \Big(E_Q[\,-X\,] - \alpha(Q) \Big) \qquad \textit{for all } X \in \mathcal{X}. \tag{16}$$

2. *The acceptance set $\mathcal{A}_\rho$ associated with ρ is weak*-, i.e., $\sigma(L^\infty(P), L^1(P))$-closed.*
3. *ρ possesses the Fatou property: If the sequence $(X_n)_{n\in\mathbb{N}} \subset \mathcal{X}$ is uniformly bounded, and X_n converges to some $X \in \mathcal{X}$ in probability, then $\rho(X) \leq \liminf_n \rho(X_n)$.*
4. *If the sequence $(X_n)_{n\in\mathbb{N}} \subset \mathcal{X}$ decreases to $X \in \mathcal{X}$, then $\rho(X_n) \to \rho(X)$.*

Proof $1 \Rightarrow 2$ holds, because ρ given by (16) is $\sigma(L^\infty(P), L^1(P))$-lower semi-continuous. For the converse implication, we can repeat the proof of Theorem 5 and apply the Hahn-Banach separation theorem in the locally convex space $\big(L^\infty(P), \sigma(L^\infty(P), L^1(P))\big)$ in order to get a negative continuous linear functional ℓ satisfying (14). By assumption, ℓ can be represented as $\ell(Z) = E[\,\varphi X\,]$ with some $\varphi \in L^1(P)$ yielding a probability measure $dQ/dP = \varphi / E[\,\varphi]$. We conclude the proof as in Theorem 5. The remaining implications follow as in [4]. □

Proposition 7 *Suppose $\rho : L^\infty(\Omega, \mathcal{F}, P) \to \mathbb{R}$ is a convex measure of risk possessing a representation of the form (16) and take $\mathcal{P}$ as in Theorem 6. Then the representation (16) holds as well in terms of the penalty function*

$$\alpha_0(Q) = \sup_{X \in L^\infty} \Big(E_Q\big[-X\big] - \rho(X) \Big) = \sup_{X \in \mathcal{A}_\rho} E_Q\big[-X\big]\,. \tag{17}$$

Moreover, it is minimal in the sense that $\alpha_0(Q) \le \alpha(Q)$ for all $Q \in \mathcal{P}$ if the representation (16) holds for $\alpha(\cdot)$. In addition,

$$\alpha_0(Q) = \sup_{X \in \mathcal{A}_\rho} E_Q\big[-X\big] = \sup_{X \in \mathcal{A}} E_Q\big[-X\big] \tag{18}$$

if ρ is defined as in (8) via a given acceptance set $\mathcal{A}$.

Proof By (11) and (12) we already know that the two terms on the righthand side of (17) coincide, and the proof of our representation theorem shows that we can take $\alpha_0(\cdot)$ as in (17) once ρ possesses of the form (16). If $\alpha(\cdot)$ is any other functional with which we can represent ρ, then it follows readily that

$$\alpha(Q) \ge E_Q\big[-X\big] - \rho(X) \qquad \text{for all } X.$$

Thus $\alpha(Q) \ge \alpha_0(Q)$ for all $Q \in \mathcal{P}$. This proves the first part of the assertion. The second part follows by recalling that $\mathcal{A} \subseteq \mathcal{A}_\rho$ and that $\varepsilon + X \in \mathcal{A}$ whenever $X \in \mathcal{A}_\rho$ and $\varepsilon > 0$. □

Remark 8 Equation (18) shows that the minmal penalty function α_0 is lower semicontinuous for the weak* topology on $\mathcal{P}$ considered as a subset of $L^1(P)$. In particular, α_0 is lower semicontinuous for the total variation distance.

Sometimes, it may be convenient to represent a convex measure of risk with a penalty function $\alpha(\cdot)$ that is not the minimal one. This case occurs, for instance, in the following situation.

Proposition 9 *Suppose that for every i in some index set I we are given a convex measure of risk ρ_i on $\mathcal{X} := L^\infty(\Omega, \mathcal{F}, P)$ with associated penalty function $\alpha_i(\cdot)$. We assume that*

$$\inf_{Q \in \mathcal{Q}} \inf_{i \in I} \alpha_i(Q) > -\infty\,. \tag{19}$$

Then

$$\rho(X) := \sup_{i \in I} \rho_i(X)\,, \qquad X \in \mathcal{X}\,,$$

is a convex measure of risk that can be represented as in (16) with the penalty function

$$\alpha(Q) := \inf_{i \in I} \alpha_i(Q)\,, \qquad Q \ll P\,.$$

Proof Clearly,

$$\rho(X) = \sup_{i \in I} \sup_{Q \ll P} \Big(E_Q[-X] - \alpha_i(Q) \Big) = \sup_{Q \ll P} \Big(E_Q[-X] - \inf_{i \in I} \alpha_i(Q) \Big)\,.$$

Hence the assertion follows. □

3 Risk measures defined in terms of shortfall risk

Suppose that $\ell : \mathbb{R} \to \mathbb{R}$ is an increasing convex loss function which is not identically constant. For a position $X \in L^\infty(\Omega, \mathcal{F}, P)$ we introduce the expected loss

$$E_P\big[\,\ell(-X)\,\big]\,.$$

If ℓ vanishes on $(-\infty, 0]$ then $E_P\big[\,\ell(-X)\,\big] = E_P\big[\,\ell(X^-)\,\big]$ may be viewed as a quantitative assessment of the *shortfall risk*. In view of examples such as Example 12, it will be convenient to formulate our results without this restriction on ℓ.

Let x_0 be an interior point in the range of ℓ. A position $X \in L^\infty(\Omega, \mathcal{F}, P)$ will be called acceptable if the expected loss is bounded by x_0. Thus, we consider the class

$$\mathcal{A} := \Big\{\, X \in L^\infty(\Omega, \mathcal{F}, P) \,\Big|\, E_P\big[\,\ell(-X)\,\big] \le x_0 \,\Big\}\,. \tag{20}$$

of acceptable positions. The set $\mathcal{A}$ satisfies the first two properties of Proposition 2 and thus defines a convex measure of risk $\rho := \rho_{\mathcal{A}}$. Since ℓ is continuous as a finitely valued convex function on $\mathbb{R}$, ρ possesses the Fatou property and, hence, a representation of the form (16). The corresponding minimal penalty function $\alpha_0(\cdot)$ can be expressed in terms of the Fenchel-Legendre transform

$$\ell^*(z) := \sup_{x \in \mathbb{R}} \big(\, zx - \ell(x) \,\big)$$

of ℓ; note that the following formula may be viewed as an extension of a classical result for Orlicz spaces; cf. [13, p. 91].

Theorem 10 *Suppose that $\mathcal{A}$ is the acceptance set given by (20). Then, for $Q \ll P$, the minimal penalty function of $\rho = \rho_{\mathcal{A}}$ is given by*

$$\alpha_0(Q) = \sup_{X \in \mathcal{A}} E_Q[\,-X\,] = \inf_{\lambda > 0} \frac{1}{\lambda}\Big(x_0 + E_P\Big[\,\ell^*\Big(\lambda \frac{dQ}{dP}\Big)\,\Big]\Big)\,. \tag{21}$$

For the convenience of the reader, we summarize below some basic properties of the functions ℓ and ℓ^*.

Lemma 11 *The functions ℓ and ℓ^* enjoy the following properties.*

1. $\ell^*(0) = -\inf_{x \in \mathbb{R}} \ell(x)$ *and* $\ell^*(z) \ge -\ell(0)$ *for all* z.
2. *The set* $N := \big\{\, z \in \mathbb{R} \,\big|\, \ell^*(z) = -\ell(0) \,\big\}$ *is non-empty,* $z_1 := \inf N \ge 0$, *and* $\ell^*(z) = \sup_{x \ge 0} \big(xz - \ell(x)\big)$ *for* $z \ge z_1$. *In particular, ℓ^* is non-decreasing on* $[z_1, \infty)$.
3. $z_0 := \inf \big\{\, z \in \mathbb{R} \,\big|\, \ell^*(z) < \infty \,\big\} \in [0, \infty)$.
4. $\dfrac{\ell^*(z)}{z} \longrightarrow \infty$ *as* $z \uparrow \infty$.

Proof Assertion 1 is trivial. As for the second claim, note that convexity of ℓ implies that the set of all z with $zx \le \ell(x) - \ell(0)$ for all $x \in \mathbb{R}$ is non-empty. For those z we clearly have $\ell^*(z) \le -\ell(0)$. On the other hand, $\ell^*(z) \ge -\ell(0)$ by 1. Next, it is clear that $z_1 \ge 0$. If $z \ge z_1$ and $x < 0$, then $xz - \ell(x) \le xz_1 - \ell(x) \le \ell^*(z_1) \le -\ell(0)$, where the latter inequality follows from the lower semicontinuity of ℓ^*. But $\ell^*(z) \ge -\ell(0)$ so that the supremum of $(xz - \ell(x))$ must occur for $x \ge 0$.

3. By 2, ℓ^* is not identically $+\infty$ and, thus, $z_0 < \infty$. On the other hand, $\ell^*(z) = +\infty$ for all $z < 0$.

4. For $z \ge z_1$, $\ell^*(z)/z = \sup_{x \ge 0} \big(x - \ell(x)/z\big)$ by 2. Hence $\ell^*(z)/z \ge x_z - 1$ if $x_z = \sup\big\{\, x \,\big|\, \ell(x) \le z \,\big\}$. Since ℓ is increasing and takes only finite values, it follows that $x_z \to \infty$ as $z \uparrow \infty$. □

Proof of Theorem 10 Fix $Q \ll P$ and let

$$\varphi := \frac{dQ}{dP}\,.$$

First, let us remark that it suffices to consider the case where $x_0 > \ell(0)$. Otherwise we can find some $a \in \mathbb{R}$ such that $\ell(-a) < x_0$, since x_0 was assumed to be an interior point of $\ell(\mathbb{R})$. Let $\widetilde{\ell}(x) := \ell(x - a)$, and

$$\widetilde{\mathcal{A}} := \Big\{\, \widetilde{X} \in L^\infty(\Omega, \mathcal{F}, P) \,\Big|\, E_P\big[\,\widetilde{\ell}(-\widetilde{X})\,\big] \le x_0 \,\Big\}\,.$$

Then $\widetilde{\mathcal{A}} = \big\{\, X - a \,|\, X \in \mathcal{A} \,\big\}$, and hence

$$\sup_{\widetilde{X} \in \widetilde{\mathcal{A}}} E_Q[\,-\widetilde{X}\,] = \sup_{X \in \mathcal{A}} E_Q[\,-X\,] + a\,. \tag{22}$$

The loss function $\widetilde{\ell}$ satisfies the assumption $\widetilde{\ell}(0) < x_0$. So if the assertion is established in this case, we find that

$$\sup_{\widetilde{X} \in \widetilde{\mathcal{A}}} E_Q[\,-\widetilde{X}\,] = \inf_{\lambda > 0} \frac{1}{\lambda}\Big(x_0 + E_P\Big[\,\widetilde{\ell}^*(\lambda\varphi)\,\Big]\Big) = \inf_{\lambda > 0} \frac{1}{\lambda}\Big(x_0 + E_P\Big[\,\ell^*(\lambda\varphi)\,\Big]\Big) + a\,,$$

where we have used that the Fenchel-Legendre transform $\widetilde{\ell}^*$ of $\widetilde{\ell}$ satisfies $\widetilde{\ell}^*(z) = \ell^*(z) + az$. Together with (22), this proves that the reduction to the case $\ell(0) < x_0$ is indeed justified.

Next, let I denote the right-continuous derivative of ℓ^*, and recall that for $x, z \in \mathbb{R}$

$$xz \le \ell(x) + \ell^*(z) \qquad \text{with equality if } x = I(z); \tag{23}$$

see, e.g., [14]. Thus, for any $\lambda > 0$ and $X \in \mathcal{A}$ we have that

$$-X\varphi = \frac{1}{\lambda}(-X)(\lambda\varphi) \le \frac{1}{\lambda}\Big(\ell(-X) + \ell^*(\lambda\varphi)\Big)\,.$$

Hence, for any $\lambda > 0$

$$\alpha_0(Q) \le \sup_{X \in \mathcal{A}} \frac{1}{\lambda}\Big(E_P\big[\,\ell(-X)\,\big] + E_P\big[\,\ell^*(\lambda\varphi)\big]\Big) \le \frac{1}{\lambda}\Big(x_0 + E_P\big[\,\ell^*(\lambda\varphi)\big]\Big)$$

Thus, it remains to prove that

$$\alpha_0(Q) \geq \inf_{\lambda>0} \frac{1}{\lambda}\Big(x_0 + E_P\big[\,\ell^*(\lambda\varphi)\big]\Big) \tag{24}$$

in case where $\alpha_0(Q) < \infty$. This will be done first under the following two extra conditions:

$$\begin{aligned} &\ell^* < \infty \text{ on } (0,\infty) \text{ and there exists } \kappa \in \mathbb{R} \\ &\qquad \text{such that } \ell(x) = \inf \ell \text{ for all } x \leq \kappa. \end{aligned} \tag{25}$$

$$I \text{ is continuous on } (0,\infty). \tag{26}$$

Note that (25) implies that $\ell^*(0) < \infty$ and that $I(0+) \geq \kappa$. Since $\ell^*(z) \geq -\ell(0) > -x_0$, it follows from (23) that

$$\lim_{z\downarrow 0} \ell(I(z)) - x_0 < \lim_{z\downarrow 0} \Big(\,\ell(I(z)) + \ell^*(z)\,\Big) = \lim_{z\downarrow 0} zI(z) = 0\,.$$

Since ℓ is non-constant, convex, increasing, and takes only finite values, we may conclude that the continuous function $I(z)$ increases to $+\infty$ as $z \uparrow \infty$, and hence so does $\ell(I(z))$. Thus, there exits some $\lambda_c > 0$ such that

$$E\Big[\,\ell\Big(I(\lambda_c\varphi)\mathrm{I}_{\{\varphi\leq c\}}\Big)\,\Big] = x_0$$

for $c > 0$ large enough. Let

$$X^c := -I(\lambda_c\varphi)\mathrm{I}_{\{\varphi\leq c\}}\,.$$

Then X^c is bounded and contained in $\mathcal{A}$. Moreover, by (23),

$$-X^c\varphi = \frac{1}{\lambda_c}(-X^c)(\lambda_c\varphi) = \frac{1}{\lambda_c}\Big(\ell\big(I(\lambda_c\varphi)\big) + \ell^*(\lambda_c\varphi)\Big)\mathrm{I}_{\{\varphi\leq c\}}\,.$$

It follows that

$$\alpha_0(Q) \geq E_Q\big[\,-X^c\,\big] = \frac{1}{\lambda_c}\Big(x_0 - \ell(0)\cdot P\big[\,\varphi > c\,\big] + E\Big[\,\ell^*(\lambda_c\varphi)\cdot \mathrm{I}_{\{\varphi\leq c\}}\,\Big]\Big)\,.$$

The right-hand side is bounded below by $(x_0 - \ell(0))/\lambda_c$. As we assumed that $\alpha_0(Q) < \infty$, we must have that the decreasing limit $\lambda_\infty = \lim_{c\uparrow\infty}\lambda_c$ is strictly positive. Thus, we obtain that

$$\alpha_0(Q) \geq \liminf_{c\uparrow\infty} E_Q\big[\,-X^c\,\big] \geq \frac{1}{\lambda_\infty}\Big(x_0 + E_P\Big[\,\ell^*(\lambda_\infty\varphi)\,\Big]\Big)\,,$$

where we have used Fatou's lemma together with the fact that $\ell^*(z) \geq -\ell(0)$. This proves (24) and, hence, (21) in the case in which I is continuous and ℓ satisfies (25).

If I is not continuous, then we can approximate the upper semicontinuous function I arbitrarily close from above with an increasing continuous function $\widetilde{I}$ such that

$$\widetilde{\ell}^*(z) := \ell^*(0) + \int_0^z \widetilde{I}(y)\,dy$$

satisfies

$$\ell^*\big(z\big) \le \widetilde{\ell}^*(z) \le \ell^*\big((1+\varepsilon)z\big) \qquad \text{for } z \ge 0.$$

Let $\widetilde{\ell} := \widetilde{\ell}^{**}$ denote the Fenchel-Legendre transform of $\widetilde{\ell}^*$. Since $\ell^{**} = \ell$, it follows that

$$\ell\Big(\frac{x}{1+\varepsilon}\Big) \le \widetilde{\ell}(x) \le \ell(x)\,.$$

Therefore,

$$\widetilde{\mathcal{A}} := \big\{\, X \in \mathcal{X} \,\big|\, E_P\big[\,\widetilde{\ell}(-X)\,\big] \le x_0 \,\big\} \subseteq \big\{\, (1+\varepsilon)X \,\big|\, X \in \mathcal{A} \,\big\} =: \mathcal{A}_\varepsilon\,.$$

Since we already know that the assertion holds for $\widetilde{\ell}$, we obtain that

$$\begin{aligned}
\inf_{\lambda>0} \frac{1}{\lambda}\Big(x_0 + E_P\Big[\,\ell^*\Big(\lambda\frac{dQ}{dP}\Big)\,\Big]\Big) &\le \inf_{\lambda>0} \frac{1}{\lambda}\Big(x_0 + E_P\Big[\,\widetilde{\ell}^*\Big(\lambda\frac{dQ}{dP}\Big)\,\Big]\Big) \\
&= \sup_{X\in\widetilde{\mathcal{A}}} E_Q[\,-X\,] \\
&\le \sup_{X\in\mathcal{A}_\varepsilon} E_Q[\,-X\,] \\
&= (1+\varepsilon)\alpha_0(Q)\,.
\end{aligned}$$

This proves (21) in case where ℓ satisfies (25).

If ℓ does not satisfy (25) we can find a sequence (ℓ_n) of convex loss functions with $\ell_n(0) = \ell(0)$ such that (25) holds for each n and such that $\ell_n \searrow \ell$ pointwise. For instance, we can take

$$\ell_n(x) := \ell(x) \vee \kappa_n + \frac{1}{n}(e^x - 1)^+\,,$$

where (κ_n) is a sequence such that $\kappa_n \le \ell(0)$ and $\kappa_N \downarrow \inf \ell$. One checks that $\ell_n^*(z) \nearrow \ell^*(z)$ for all z and that the corresponding minimal penalty functions α_0^n increase to α_0. From the preceding, we know that

$$\infty > \alpha_0(Q) \ge \alpha_0^n(Q) \ge \frac{1}{\lambda_n^\varepsilon}\Big(x_0 + E_P\Big[\,\ell_n^*(\lambda_n^\varepsilon\varphi)\,\Big]\Big) - \varepsilon$$

for certain $\lambda_n^\varepsilon \in (0,\infty)$. The fourth assertion of Lemma 11 implies that $\sup_n \lambda_n^\varepsilon < \infty$ if $\alpha_0(Q) < \infty$. Moreover, we have for large enough n that $\inf_z \ell_n^*(z) = -\ell_n(0) = -\ell(0) > x_0$. Hence, $\liminf_n \lambda_n^\varepsilon$ must be bounded away from 0 if $\alpha_0(Q) < \infty$. Hence, we may assume that λ_n^ε converges towards some $\lambda^\varepsilon \in (0,\infty)$. Then

$$\alpha_0(Q) + \varepsilon \ge \liminf_{n\uparrow\infty} \frac{1}{\lambda_n^\varepsilon}\Big(x_0 + E_P\Big[\,\ell_n^*(\lambda_n^\varepsilon\varphi)\,\Big]\Big) \ge \frac{1}{\lambda^\varepsilon}\Big(x_0 + E_P\Big[\,\ell^*(\lambda^\varepsilon\varphi)\,\Big]\Big)\,,$$

where we used again that $\ell_n^*(z) \ge -\ell(0)$, uniformly in n and z. Letting $\varepsilon \downarrow 0$ completes the proof of the theorem. □

Example 12 For an exponential loss function, the penalty function can be described in terms of relative entropy. Let us take $\ell(x) = e^x$ and $x_0 = 1$ so that

$$\rho(X) = \inf\left\{ m \in \mathbb{R} \,\middle|\, E_P\big[\, e^{-m-X} \,\big] \le 1 \right\} = \log E_P\big[\, e^{-X} \,\big]\,.$$

Then (17) becomes the well-known variational formula for the relative entropy, namely

$$\alpha(Q) = \sup_{X \in L^\infty(\Omega,\mathcal{F},P)} \Big(E_Q\big[\, -X \,\big] - \log E_P\big[\, e^{-X} \,\big] \Big) = H(Q|P)\,,$$

where the relative entropy of Q with respect to P is defined as

$$H(Q|P) := \begin{cases} \displaystyle\int \frac{dQ}{dP} \log \frac{dQ}{dP}\, dP & \text{if} Q \ll P, \\ +\infty & \text{otherwise.} \end{cases}$$

Example 13 Take in (20)

$$\ell(x) := \begin{cases} \frac{1}{p} x^p & \text{if } x \ge 0, \\ 0 & \text{otherwise,} \end{cases}$$

where $p > 1$. Then $I(z) = z^{1/(p-1)}$, and with $q = p/(p-1)$

$$\ell^*(z) := \begin{cases} \frac{1}{q} z^q & \text{if } x \ge 0, \\ +\infty & \text{otherwise.} \end{cases}$$

Thus, for $x_0 > 0$ the conditions of Theorem 10 are fulfilled. Thus, if $Q \ll P$ with $\varphi := dQ/dP \in L^q(\Omega, \mathcal{F}, P)$, then the infimum in (21) is attained for

$$\lambda_Q = \Big(\frac{x_0 p}{E_P[\, \varphi^q \,]} \Big)^{1/q}.$$

Hence we can identify $\alpha(Q)$ for any $Q \ll P$ as

$$\alpha(Q) = (p x_0)^{1/p} \cdot E_P\Big[\Big(\frac{dQ}{dP} \Big)^q \Big]^{1/q}.$$

Together with Proposition 9, Theorem 10 yields the following result for risk measures which are defined in terms of a robust notion of bounded shortfall risk.

Corollary 14 *Suppose that $\mathcal{Q}$ is a family of equivalent probability measures, and that ℓ, ℓ^*, and x_0 are as in Theorem 10. On $\mathcal{X} := L^\infty(\Omega, \mathcal{F}, P)$, for any $P \in \mathcal{Q}$, let*

$$\mathcal{A} := \left\{ X \in \mathcal{X} \,\middle|\, E_P\big[\, \ell(-X) \,\big] \le x_0 \text{ for all } P \in \mathcal{Q} \right\}. \tag{27}$$

Then the corresponding convex risk measure can be represented in terms of the penalty function

$$\alpha(Q) = \inf_{\lambda > 0} \frac{1}{\lambda} \Big(x_0 + \inf_{P \in \mathcal{Q}} E_P\Big[\, \ell^*\Big(\lambda \frac{dQ}{dP} \Big) \,\Big] \Big)\,.$$

Example 15 In the situation of Example 12, the corresponding robust problem as in Corollary 14 leads to the following entropy minimization problem: For a given Q and a set $\mathcal{Q}$ of equivalent probabilities, find

$$\inf_{P\in\mathcal{Q}} H(Q|P)\,.$$

Note that this problem is different from the standard problem of minimizing $H(Q|P)$ with respect to the first variable as it appears, e.g., in the theory of large deviations.

4 Risk measures arising in a financial market with convex constraints

We consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0,\dots,T}, P)$ and a market where one bond and d risky assets are traded. The price of the bond will be assumed to be normalized to 1, and the (correspondingly discounted) price process of the risky assets is denoted by $X_t = (X^1, \dots, X^d_t)$. We will assume that

$$X^i_t \geq 0 \qquad \text{for } i = 1, \dots, d \text{ and } t = 0, \dots, T. \tag{28}$$

Any d-dimensional predictable process ξ can be regarded as a self-financing trading strategy; ξ^i_t is the number of shares held of the i^{th} asset during the trading period $t-1 \rightsquigarrow t$, and

$$V_t = V_0 + \sum_{k=1}^{t} \xi_k \cdot (X_k - X_{k-1})$$

is the associated value process for an initial endowment V_0.

Now consider a financial position $Z \in L^\infty(P)$. Z can be interpreted as "riskless" if $Z \geq 0$ or, more generally, if the "risky part" of Z can be hedged at no additional cost. The latter means that we can find a suitable hedging portfolio ξ such that

$$Z + \sum_{t=1}^{T} \xi_t \cdot (X_t - X_{t-1}) \geq 0 \qquad P\text{-a.s.} \tag{29}$$

Note that (29) is only possible if ξ is *admissible* in the sense that there is a constant $c = c(\xi)$ such that the associated gains process satisfies

$$G_T(\xi) := \sum_{t=1}^{T} \xi_t \cdot (X_t - X_{t-1}) \geq -c \qquad P\text{-a.s.}, \tag{30}$$

because the position Z is bounded.

If we think of Z being the value of a large portfolio, we should avoid to run into the regime of illiquidity when hedging Z. For instance, one may want to impose individual limits on the number of shares held of the risky assets. In this case, each component ξ^i is only allowed to take values in some interval $[a_i, b_i]$ with $a_i \leq 0 \leq b_i$. Such constraints on the hedging portfolio were first suggested by Cvitanic and Karatzas [2], [3]. Here we will work with the more general framework introduced in the continuous-time setting by Föllmer and Kramkov [5].

Thus, let $\mathcal{S}$ be a set of admissible trading strategies such that

– $0 \in \mathcal{S}$.
– $\mathcal{X}$ is *predictably convex*: If $\xi^1, \xi^2 \in \mathcal{S}$ and h is a predictable process with $0 \le h \le 1$, then also $\xi^h \in \mathcal{S}$ where $\xi^h_t := h_t\xi^1_t + (1-h_t)\xi^2_t$.

In a first step, we define the non-empty set

$$\mathcal{A} := \Big\{ Z \in L^\infty(P) \;\Big|\; \text{there exists } \xi \in \mathcal{S} \text{ with } Z + \sum_{t=1}^{T} \xi_t \cdot (X_t - X_{t-1}) \ge 0 \; P\text{-a.s.} \Big\}, \quad (31)$$

of acceptable positions which can be hedged with strategies in $\mathcal{S}$ at no cost. By Proposition 4, $\mathcal{A}$ induces the convex measure of risk

$$\rho(Z) := \rho_{\mathcal{A}}(Z) = \inf \Big\{ m \in \mathbb{R} \;\Big|\; m + Z \in \mathcal{A} \Big\} \quad (32)$$

provided that

$$\rho_{\mathcal{A}}(0) > -\infty\,. \quad (33)$$

Note that (33) holds, in particular, if $\mathcal{S}$ does not contain arbitrage opportunities. We will assume (33) throughout this section.

The following questions arise:

– When does ρ permit a representation of the form (16)?
– If so, can one identify the minimal penalty function α_0?

Let us first consider the second question.

Proposition 16 *Suppose the convex measure of risk $\rho_{\mathcal{A}}$ induced by the set $\mathcal{A}$ of (31) possesses a representation of the form (16). Then the minimal penalty function $\alpha_0(\cdot)$ in the representation (16) is given by*

$$\alpha_0(Q) = E_Q\big[\, A^Q_T \,\big] \qquad \textit{for } Q \ll P. \quad (34)$$

Here A^Q is the predictable increasing process defined by

$$A^Q_0 := 0\,, \qquad A^Q_{t+1} - A^Q_t := \operatorname*{ess\,sup}_{\xi \in \mathcal{S}} \Big[\xi_{t+1} \cdot \Big(E_Q\big[\, X_{t+1} \,\big|\, \mathcal{F}_t \,\big] - X_t \Big) \Big]. \quad (35)$$

Proof First note that $E_Q\big[\, \xi_t \cdot (X_t - X_{t-1}) \,\big|\, \mathcal{F}_{t-1} \,\big]$ is well-defined and satisfies

$$E_Q\big[\, \xi_t \cdot (X_t - X_{t-1}) \,\big|\, \mathcal{F}_{t-1} \,\big] = \xi_t \cdot \Big(E_Q\big[\, X_t \,\big|\, \mathcal{F}_{t-1} \,\big] - X_{t-1} \Big)$$

for every $\xi \in \mathcal{S}$. To see this, observe first that by predictable convexity also $\xi^{(t)} \in \mathcal{S}$, where

$$\xi^{(t)}_s := \begin{cases} \xi_t & \text{if } s = t, \\ 0 & \text{otherwise.} \end{cases}$$

By assumption every element in $\mathcal{S}$ is admissible in the sense of (30), and thus there is some constant c with $\xi_t \cdot (X_t - X_{t-1}) = G_T(\xi^{(t)}) \ge -c$ P-a.s. Using our Assumption (28) that prices are non-negative, the claim follows.

Next, if $Z \in \mathcal{A}$ there exists $\xi \in \mathcal{S}$ with $-Z \leq \sum_{t=1}^{T} \xi_t \cdot (X_t - X_{t-1})$ P-a.s. By using the above identity for $\xi_t \cdot \big(E_Q\big[\, X_t \,\big|\, \mathcal{F}_{t-1} \,\big] - X_{t-1}\big)$, we obtain that for $Q \ll P$

$$E_Q[\,-Z\,] \leq E_Q\Big[\, \sum_{t=1}^{T} \xi_t \cdot (X_t - X_{t-1}) \Big] \leq E_Q\big[\, A_T^Q \,\big]\,. \tag{36}$$

Hence, we conclude from Proposition 7 that

$$\alpha_0(Q) = \sup_{Z \in \mathcal{A}} E_Q[\,-Z\,] \leq E_Q\big[\, A_T^Q \,\big]\,.$$

Now we turn to the proof of the converse inequality. To this end, we show first that

$$\Psi := \left\{ \sum_{t=1}^{T} E_Q\big[\, \xi_t \cdot (X_t - X_{t-1}) \,\big|\, \mathcal{F}_{t-1} \,\big] \;\middle|\; \xi \in \mathcal{S} \right\}$$

is directed upwards in the sense that for $\psi_1, \psi_2 \in \Psi$ there is $\psi_3 \in \Psi$ with $\psi_3 \geq \psi_1 \vee \psi_2$. For $\xi, \widetilde{\xi} \in \mathcal{S}$ let

$$A_t := \left\{ E_Q\big[\, \xi_t \cdot (X_t - X_{t-1}) \,\big|\, \mathcal{F}_{t-1} \,\big] > E_Q\big[\, \widetilde{\xi}_t \cdot (X_t - X_{t-1}) \,\big|\, \mathcal{F}_{t-1} \,\big] \right\},$$

and define $\xi' \in \mathcal{S}$ by

$$\xi'_t := \xi_t \mathrm{I}_{A_t} + \widetilde{\xi}_t \mathrm{I}_{A_t^c}\,.$$

Then clearly

$$\begin{aligned} &E_Q\big[\, \xi'_t \cdot (X_t - X_{t-1}) \,\big|\, \mathcal{F}_{t-1} \,\big] \\ &\quad = E_Q\big[\, \xi_t \cdot (X_t - X_{t-1}) \,\big|\, \mathcal{F}_{t-1} \,\big] \vee E_Q\big[\, \widetilde{\xi}_t \cdot (X_t - X_{t-1}) \,\big|\, \mathcal{F}_{t-1} \,\big]\,, \end{aligned}$$

and therefore Ψ is directed upwards. It follows that $A_T^Q = \operatorname{ess\,sup} \Psi$ is the limit of an increasing sequence in Ψ. Hence,

$$\begin{aligned} E_Q\big[\, A_T^Q \,\big] &= \sup_{\xi \in \mathcal{S}} E_Q\Big[\sum_{t=1}^{T} E_Q\big[\, \xi_t \cdot (X_t - X_{t-1}) \,\big|\, \mathcal{F}_{t-1} \,\big] \Big] \\ &= \sup_{\xi \in \mathcal{S}} \sup_{k \in \mathbb{N}} E_Q\big[\, G_T(\xi) \wedge k \,\big]\,. \end{aligned}$$

Admissibility yields that $-(G_T(\xi) \wedge k) \in \mathcal{A} \subseteq \mathcal{A}_\rho$, and thus

$$E_Q\big[\, A_T^Q \,\big] = \sup_{\xi \in \mathcal{S}} \sup_{k \in \mathbb{N}} E_Q\big[\, G_T(\xi) \wedge k \,\big] \leq \alpha_0(Q)\,.$$

This concludes the proof. □

Let us now turn to the question of the existence of a representation (16). Let $\mathcal{P}_{\mathcal{S}}$ denote the class of all probability measures $\widetilde{P}$ that satisfy the following conditions

- $\widetilde{P} \approx P$,
- $X_t \in L^1(\widetilde{P})$ for all t, (37)
- the value process of every strategy in $\mathcal{S}$ is a supermartingale under $\widetilde{P}$.

Theorem 17 *Assume that $\mathcal{S}$ is closed under convergence in probability, and let ρ denote the convex measure of risk on $\mathcal{X} = L^\infty(P)$ arising from the set $\mathcal{A}$ of (31). Then the following conditions are equivalent.*

1. *ρ is relevant in the sense that $\rho(X) > 0$ whenever $X \in \mathcal{X}$ is non-positive and satisfies $P[\, X < 0\,] > 0$.*
2. *There are no arbitrage opportunities in $\mathcal{S}$.*
3. *The set $\mathcal{P}_{\mathcal{S}}$ is non-empty.*

If one of these conditions is fulfilled, ρ possesses a representation of the form (16) where one can take $\alpha(Q)$ as in (34) for $Q \approx P$ and $\alpha(Q) = \infty$ if $Q \not\approx P$.

That the relevance of ρ is equivalent to the absence of arbitrage follows immediately from the respective definitions. The proof of the remaining assertions uses standard techniques for proving the fundamental theorem of asset pricing and the superhedging duality theorem; for the details we refer to [8].

Remark 18 In a continuous-time financial market model where the price process X follows a special semimartingale under P, one can similarly define a predictably convex set $\mathcal{S}$ of admissible integrands and a corresponding convex measure of risk ρ. If one assumes in addition that the set $\left\{ \int \xi\, dX \,\middle|\, \xi \in \mathcal{S} \right\}$ is closed in the seminartingale or Émery topology, the optional decomposition theorem of [5] will imply a representation (16) of ρ. The penalty function $\alpha(Q)$ can be described as $\alpha(Q) = E_Q\big[\, A_T^Q \,\big]$ provided that Q satisfies the following three conditions are fulfilled. Q is equivalent to P, every process $\int \xi\, dX$ with $\xi \in \mathcal{S}$ is a special semimartingale under Q, and Q admits the upper variation process A^Q for the set $\left\{ \int \xi\, dX \,\middle|\, \xi \in \mathcal{S} \right\}$. One can take $\alpha(Q) = \infty$ for measures Q which do not fulfill one of these conditions.

In a second step, we combine the results of this section with those of Sect. 3. If ℓ is a convex loss function and x_0 is an interior point in the range of ℓ, one can call a position $X \in \mathcal{X} = L^\infty(P)$ acceptable if there exists $\xi \in \mathcal{S}$ such that the expected loss of $X + G_T(\xi)$ is bounded by x_0, i.e.,

$$E_P\Big[\, \ell\big(-X - G_T(\xi)\big)\,\Big] \leq x_0\,. \tag{38}$$

The risk measure arising from this class of acceptable positions is closely related to the problem of efficient hedging or the problem of utility maximization; see, e.g., [6], [15]. Here we go one step further and replace the risk measure associated with ℓ by a general risk measure ρ^0 which possesses a representation (16). A position $X \in \mathcal{X}$ will be acceptable if there exists $\xi \in \mathcal{S}$ such that $X + G_T(\xi)$ is acceptable for ρ^0 in the sense that there exists $Y \in \mathcal{A}^0 := \mathcal{A}_{\rho^0}$ such that $X + G_T(\xi) \geq Y$ (note that $X + G_T(\xi)$ needs not be bounded):

$$\overline{\mathcal{A}} := \left\{ X \in \mathcal{X} \,\middle|\, \exists\, \xi \in \mathcal{S} \text{ such that } X + G_T(\xi) \geq Y \text{ for some } Y \in \mathcal{A}^0 \right\}. \tag{39}$$

Let $\overline{\rho} := \rho_{\overline{\mathcal{A}}}$. According to Proposition 4, $\overline{\rho}$ is a convex measure of risk provided that

$$\overline{\rho}(0) > -\infty\,. \tag{40}$$

Note that this condition (40) implies our Assumption (33).

Let us first consider the special case where $\mathcal{A}^0$ is defined in terms of a convex loss function, and let us show that $\overline{\rho}$ coincides indeed with the risk measure which arises from the acceptability criterion (38). For this identity, the condition (40) is not needed.

Proposition 19 *Suppose that the acceptance set $\mathcal{A}^0$ is of the form (20) for a convex loss function ℓ and an interior point x_0 in the range of ℓ. Then*

$$\overline{\rho}(X) = \inf \Big\{ m \in \mathbb{R} \,\Big|\, \exists\, \xi \in \mathcal{S} \text{ such that } E_P\big[\, \ell\big(-m - X - G_T(\xi)\big)\,\big] \le x_0 \Big\}. \tag{41}$$

Proof Denote for the moment the right-hand side of (41) by $\widehat{\rho}(X)$. Clearly, $\widehat{\rho}$ is translation invariant, and every $X \in \overline{\mathcal{A}}$ satisfies $\widehat{\rho}(X) \le 0$. Thus, we obtain $\overline{\rho} \ge \widehat{\rho}$. Conversely, if $X \in \mathcal{X}$ is arbitrary and $m > \widehat{\rho}(X)$, then there exists $\xi \in \mathcal{S}$ such that

$$E_P\big[\, \ell\big(-m - X - G_T(\xi)\big)\,\big] < x_0\,.$$

If $k > 0$ is large enough, then

$$E_P\big[\, \ell\big(-m - X - G_T(\xi) \wedge k\big)\,\big] \le x_0\,.$$

It follows that $Y := m + X + G_T(\xi) \wedge k \in \mathcal{A}^0$, which in turn implies that $m + X \in \overline{\mathcal{A}}$ and $\overline{\rho}(X) \le m$. □

For a general risk measure ρ^0, the following proposition shows how the resulting risk measure $\overline{\rho}$ depends both on ρ^0 and on the structure of the financial market.

Proposition 20 *Assume (40) and let $\rho_{\mathcal{A}}$ and $\overline{\rho}$ be the convex measures of risk associated with the acceptance sets $\mathcal{A}$ and $\overline{\mathcal{A}}$ of (31) and (39), respectively. If both $\rho_{\mathcal{A}}$ and $\overline{\rho}$ possess representations (16), then the minimal penalty function $\overline{\alpha}_0$ arising in the representation (16) of $\overline{\rho}$ is given by*

$$\overline{\alpha}_0(Q) = \alpha_0^0(Q) + \alpha_0^1(Q)\,, \qquad Q \ll P\,, \tag{42}$$

where α_0^0 is the minimal penalty function for ρ^0 and α_0^1 is given by (34).

Proof Let $\mathcal{A}$ be the acceptance set of (31). It is easy to prove that

$$\overline{\mathcal{A}} = \Big\{ X^0 + X^1 \,\Big|\, X^0 \in \mathcal{A}^0\,,\ X^1 \in \mathcal{A} \Big\}.$$

For instance, if $X \in \overline{\mathcal{A}}$, then there exists $Y \in \mathcal{A}^0$ and $\xi \in \mathcal{S}$ such that $X + G_T(\xi) \ge Y$. Hence, X^0 is the sum of $X - Y \in \mathcal{A}$ and of $Y \in \mathcal{A}^0$.

Now it follows from (18) that

$$\overline{\alpha}_0(Q) = \sup_{X \in \overline{\mathcal{A}}} E_Q\big[\,-X\,\big] = \sup_{X^0 \in \mathcal{A}^0} \sup_{X^1 \in \mathcal{A}^1} E_Q\big[\,-X^0 - X^1\,\big] = \alpha_0^0(Q) + \alpha_0^1(Q)\,.$$

This proves the assertion. □

References

1. Artzner, P., Delbaen, F., Eber, J.-M., Heath, D.: Coherent measures of risk. Math. Finance **9** (3), 203–228 (1999)
2. Cvitanic, J., Karatzas, I.: Convex duality in constrained portfolio optimization. Ann. Appl. Probab. **2** (4), 767–818 (1992)
3. Cvitanic, J., Karatzas, I.: Hedging contingent claims with constrained portfolios. Ann. Appl. Probab. **3** (3), 652–681 (1993)
4. Delbaen, F.: Coherent measures of risk on general probability spaces. Preprint ETH Zürich 2000
5. Föllmer, H., Kramkov, D.: Optional decompositions under constraints. Probab. Theory Relat. Fields **109** (1), 1–25 (1997)
6. Föllmer, H., Leukert, P.: Efficient hedging: cost versus shortfall risk. Finance Stochast. 4, 117–146 (2000)
7. Föllmer, H., Schied, A.: Robust representation of convex measures of risk. In: Sandmann, K. Schönbucher, P. (eds.) Advances in finance and stochastics. Berlin Heidelberg New York: Springer 2002
8. Föllmer, H., Schied, A.: Stochastic finance: an introduction in discrete time. Berlin: de Gruyter Studies in Mathematics (to appear)
9. Heath, D.: Back to the future. Paris: Plenary lecture at the First World Congress of the Bachelier Society 2000
10. Heath, D., Ku, H.: Market equilibrium with coherent measures of risk. Preprint, Pittsburgh, PA: Carnegie Mellon University 2001
11. Huber, P.: Robust statistics. New York: Wiley 1981
12. Jaschke, S., Küchler, U.: Coherent risk measures and good deal bounds. Finance Stochast. **5**, 181–200 (2001)
13. Krasnosel'skii, M., Rutickii, Ya.: Convex functions and Orlicz spaces. Gordon and Breach Science Publishers 1961
14. Rockafellar, R. T.: Convex analysis. Princeton, N. J.: Princeton University Press 1970
15. Schachermayer, W.: Optimal investment in incomplete financial markets. Proceedings of the first World Congress of the Bachelier Society, Paris 2000. Berlin Heidelberg New York: Springer 2001